

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: March 3, 2006, 19:01:52 ; Search time 188 Seconds  
(without alignments)  
539.875 Million cell updates/sec

Title: US-10-527-771-10  
Perfect score: 1253  
Sequence: 1 MSAVVAVLALFSYAEG.....DNLCDTRDASVRHQCCASP 231

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_geneseq\_21:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1253	100.0	231	8	ADK13847
2	636.5	50.8	236	8	ADK13845
3	277	22.1	424	2	AAW04321
4	277	22.1	424	9	ADY37246
5	274	21.9	425	9	ADY37295
6	263	21.0	424	9	ADY37230
7	262.5	20.9	422	9	ADY37284
8	216.5	17.3	473	5	ABG31109
9	211	16.8	425	5	ABG31108
10	191	15.2	424	9	ADY37254
11	186.5	14.9	451	9	ADY37256
12	166	13.2	217	9	ADY37286
13	163	13.0	440	2	AAW04322
14	160	12.8	216	9	ADY37296
15	154	12.3	207	9	ADY37297
16	153.5	12.3	218	2	AAW04323
17	152.5	12.2	218	2	AAW04322
18	147.5	11.8	218	9	ADY37248
19	137	10.9	227	8	ADV24015
20	137	10.9	229	8	ADV24017
21	137	10.9	238	2	AAV23611
22	137	10.9	238	8	ADV24007
23	136.5	10.9	248	2	AAV23609
24	136.5	10.9	248	8	ADV24005

25	135	10.8	200	9	ADY37250	Ady37250 Hookworm
26	132.5	10.6	220	2	AAW11485	Aaw11485 O. volvuli
27	132.5	10.6	248	2	AAW11484	Aaw11484 O. volvuli
28	131	10.5	454	2	AAW26785	Aaw26785 Sequence
29	130.5	10.4	223	4	AAW13066	Aaw13066 Onchocerc
30	130	10.4	232	8	ADV24013	Adv24013 Ancylost
31	126.5	10.1	253	4	AAW13067	Aaw13067 Onchocerc
32	126	10.1	270	2	AAV23608	Aay23608 Canine ho
33	126	10.1	270	8	ADV24004	Adv24004 Ancylost
34	124.5	9.9	303	6	ABP96844	Abp96844 Human PR
35	124.5	9.9	371	6	ABU11626	Abu11626 Human MDD
36	124.5	9.9	395	4	ABG27603	Abg27603 Nove1 hum
37	124.5	9.9	395	8	ADG98739	Adg98739 Protein f
38	124.5	9.9	437	8	ABW84725	Abw84725 Human dia
39	124.5	9.9	437	8	ABW84726	Abw84726 Human dia
40	123.5	9.9	222	2	AAW37166	Aaw37166 Haemonchu
41	123	9.8	247	4	ABB63677	Abb63677 Drosophil
42	123	9.8	415	7	ADG95552	Adg95552 Human NOV
43	122	9.7	255	5	AAW18962	Aaw18962 Mouse tes
44	121	9.7	207	5	ABG31111	Abg31111 C. elegan
45	121	9.7	308	4	ABB64393	Abb64393 Drosophil

ALIGNMENTS

RESULT 1	
ADK13847	ADK13847 standard; protein; 231 AA.
ID	
XX	
AC	ADK13847;
XX	
DT	03-JUN-2004 (first entry)
XX	
DE	Ostertagia ostertagi immunogenic protein seqid 10.
XX	
KW	antiparasitic; gene therapy; vaccine; nematode; worm;
KW	Ostertagia ostertagi protein; immunogenic protein.
XX	
OS	Ostertagia ostertagi.
XX	
PN	US2004052817-A1.
XX	
PD	18-MAR-2004.
XX	
PF	13-SEP-2002; 2002US-00243319.
XX	
PR	13-SEP-2002; 2002US-00243319.
XX	
PA	(GELD/) GELDHOF P.
PA	(VERC/) VERCAUTEREN I.
PA	(DMAE/) DE MAERE V.
PA	(CLAER/) CLAEREBOUT E.
PA	(VERC/) VERCRUYSSSE J.
XX	
PI	Geldhof P, Vercauteren I, De Maere V, Claerebout E, Vercruyssse J;
XX	
DR	WPI; 2004-247704/23.
DR	N-PSDB; ADK13846.
XX	
PT	New Ostertagia ostertagi proteins and nucleic acid sequences, useful for
PT	diagnosing, preventing or treating infections caused by Ostertagia
PT	ostertagi in animals.
XX	
PS	Claim 16; SEQ ID NO 10; 36pp; English.
XX	
CC	The invention describes a nucleic acid sequence (I) encoding an
CC	Ostertagia ostertagi protein, or part of the nucleic acid sequence that
CC	encodes an immunogenic fragment of the protein, where the nucleic acid
CC	sequence or its part has at least 85% homology with the nucleic acid
CC	sequence of the O. ostertagi gene having any of the 7 fully defined
CC	sequences of 306-1761 base pairs, as given in the specification. The
CC	composition and methods are useful for diagnosing, preventing or treating

CC infections caused by O. ostertagi in animals. This is the amino acid  
sequence of a novel immunogenic Ostertagia ostertagi protein.

XX  
SQ Sequence 231 AA;

Query Match 100.0%; Score 1253; DB 8; Length 231;  
Best Local Similarity 100.0%; Pred. No. 1.4e-118;  
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAVVVAVLALPSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMTVN 60  
Db 1 MSAVVVAVLALPSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMTVN 60  
QY 61 ADAVILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAQNIARWLYFKDSEETV 120  
Db 61 ADAVILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAQNIARWLYFKDSEETV 120  
QY 121 LQOVSWYVWSASLGFMTGKTKLDQFANQWAEPLANIANYNRNVKVGCAHKICPAQNMVWSC 180  
Db 121 LQOVSWYVWSASLGFMTGKTKLDQFANQWAEPLANIANYNRNVKVGCAHKICPAQNMVWSC 180  
QY 181 VTGSPKLAPEVNIWQEGKACVCDARPDSPCCDNLCTTRDAASVRHQCASP 231  
Db 181 VTGSPKLAPEVNIWQEGKACVCDARPDSPCCDNLCTTRDAASVRHQCASP 231

RESULT 2

ADK13845 standard; protein; 236 AA.

XX AC ADK13845;  
XX DT 03-JUN-2004 (first entry)  
XX DE Ostertagia ostertagi immunogenic protein seqid 8.  
XX KW antiparasitic; gene therapy; vaccine; nematode; worm;  
XX KW Ostertagia ostertagi protein; immunogenic protein.  
OS Ostertagia ostertagi.

XX Key Location/Qualifiers  
FH Misc-difference 37 /label= Lys, Gln  
FT /note= "Encoded by MAA"  
FT Misc-difference 64 /note= "Encoded by CAR"  
FT Misc-difference 111 /note= "Encoded by CAR"  
FT Misc-difference 121 /label= Asn, Thr  
FT /note= "Encoded by AMT"

XX PN US2004052817-A1.

XX PD 18-MAR-2004.

XX PF 13-SEP-2002; 2002US-00243319.

XX PR 13-SEP-2002; 2002US-00243319.

XX PA (GELD/) GELDHOF P.  
PA (VERC/) VERCAUTEREN I.  
PA (DMAE/) DE MAERE V.  
PA (CIAE/) CLAEREBOUT E.  
PA (VERC/) VERCRUYSSSE J.

XX PI Geldhof P, Vercauteren I, De Maere V, Claerebout E, Vercruyssse J;  
XX WPI; 2004-247704/23.  
XX DR N-PSDB; ADK13844.

XX PT New Ostertagia ostertagi proteins and nucleic acid sequences, useful for

PT diagnosing, preventing or treating infections caused by Ostertagia  
PT ostertagi in animals.

XX  
PS Claim 15; SEQ ID NO 8; 36pp; English.

XX The invention describes a nucleic acid sequence (I) encoding an  
CC Ostertagia ostertagi protein, or part of the nucleic acid sequence that  
CC encodes an immunogenic fragment of the protein, where the nucleic acid  
CC sequence of its part has at least 85% homology with the nucleic acid  
CC sequence of the O. ostertagi gene having any of the 7 fully defined  
CC sequences of 306-1761 base pairs, as given in the specification. The  
CC composition and methods are useful for diagnosing, preventing or treating  
CC infections caused by O. ostertagi in animals. This is the amino acid  
CC sequence of a novel immunogenic Ostertagia ostertagi protein.

XX  
SQ Sequence 236 AA;

Query Match 50.8%; Score 636.5; DB 8; Length 236;  
Best Local Similarity 52.5%; Pred. No. 6.1e-56;  
Matches 125; Conservative 35; Mismatches 67; Indels 11; Gaps 4;

QY 1 MSAVVVAV--LLALPSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMT 58  
Db 1 MQALIGIALYLVVTSNTEAGFCCPADLNGTDEARXIFLDFHNOVRDIAGASPLNLUT 60  
QY 59 VNADAV----ILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAQNIARWLYFKD 114  
Db 61 ---GAVQMRNVILGPAQNMKYRMDWDCNLEAKAKAMIWPCCTTPIDITISIPQNLAQMLLFQ 117  
QY 115 SEETVLQOVSWYVWSASLGFMTGKTKLDQFANQWAEPLANIANYNRNVKVGCAHKIC--PA 172  
Db 118 SQEXEVLTTQPMWSWTYASLRNLQPTDEANIYNNQIRPLSNIANWQNLKVGCAHKVCKEPT 177  
QY 173 QQNMVWSCVYSPKLAPEVNIWQEGKACVCDARPDSPCCDNLCTTRDAASVRHQCAS 230  
Db 178 GTNMVWSCAYGGEVLQDNEVWMDKGPCTCMCNAYPNSFCCNNLCTTIAATLTRKQPKS 235

RESULT 3

AAW04321 standard; protein; 424 AA.

XX AC AAW04321;  
XX DT 17-JAN-1997 (first entry)  
XX DE Ancylostoma secreted protein ASP-1 (pro-form).  
XX KW Ancylostoma secreted protein; ASP-1; hookworm; vaccine.  
OS Ancylostoma caninum.

XX Key Location/Qualifiers  
FH Peptide 1..18 /label= sig\_peptide  
FT Protein 19..424 /label= Mat\_protein

XX PN WO9632479-A1.

XX PD 17-OCT-1996.

XX PF 10-APR-1996; 96WO-US004821.

XX PR 10-APR-1995; 95US-00419414.

XX PA (UYVA ) UNIV YALE.

XX PI Hawdon JM, Hotez PJ, Jones BF;  
XX WPI; 1996-477130/47.  
XX DR N-PSDB; AAT38466.





Query Match	21.9%;	Score 274;	DB 9;	Length 425;
Best Local Similarity	30.3%;	Pred. No. 8.6e-19;		
Matches	74;	Conservative	37;	Mismatches 97; Indels 36; Gaps 11
QY	2	SAAVVAVVLAL	-----FSVAEAGFCPCPNLSQSDSARQIFLDFHNDVRRNIALG-----	N 52
DB	3	SSVVVISVISTIAFCDSAPARASFGCSNN-GITDSDRQAFLDFHNNARRRVAQVEDNKS		61
QY	53	GLINWTVNADAVILGPAQNNYKVDWDCNLEEVAAQQAQIAPCNDPLPINTSLAQNIRWLY-		111
DB	62	GKLN-----PAKNNYKLEWDCMEQQLQDAIGSCPGGSAGIQGFSQNVMSMSNS		110

QY	112	--FKDSEETVLCQVSWYVWSASLGFMKGTCLD-QFANQWAEPLANIANYNRKYCAHK	168
Db	111	GGFPNSSEK-IESITLSGWSGAK--NNGVGS DNK YTGGLYAFSNMVFSETTKIGCAYK	166
QY	169	ICPAQGNMVGSCVYGPSKLA PNEVIWQEGKACV----CDARPDSPCCDNLCDT-RDAASV	223
Db	167	VCGTK--MATSCITNGIGYITNAPMWTGQACTGADGSTYKNSGCEDSLCTKGADVPET	224
QY	224	RHQC	227
Db	225	NQC	228
RESULT 6			
ID	ADY37230	standard; protein; 424 AA.	
XX	AC	ADY37230;	
XX	DT	05-MAY-2005 (first entry)	
XX	DE	Hookworm antigen #1.	
XX	KM	antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;	
KM	KM	virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;	
KM	KM	hookworm; immune response; HIV infection; tuberculosis; malaria; measles;	
KM	KM	tetanus; diphtheria; pertussis; polio; gene; ds.	
XX	OS	Necator americanus.	
XX	PN	US2005042232-A1.	
XX	PD	24-FEB-2005.	
XX	PF	16-APR-2004; 2004US-00825692.	
XX	PR	17-OCT-2001; 2001US-0329533P.	
PR	PR	23-NOV-2001; 2001US-0332007P.	
PR	PR	26-APR-2002; 2002US-0375404P.	
PR	PR	17-OCT-2002; 2002WO-US033106.	
PR	PR	26-SEP-2003; 2003US-0505848P.	
XX	PA	(HOTE/) HOTEZ P.	
PA	PA	(ASHC/) ASHCOM J.	
PA	PA	(BDAM/) BDAMCHIAN M.	
PA	PA	(ZHAN/) ZHAN B.	
PA	PA	(WANG/) WANG Y.	
PA	PA	(HAWD/) HAWDON J.	
PA	PA	(LOUK/) LOUKAS A.	
PA	PA	(WILL/) WILLIAMSON A.	
PA	PA	(JONE/) JONES B.	
PA	PA	(BETH/) BETHONY J.	
PA	PA	(GOUD/) GOUD G.	
PA	PA	(BOT/) BOTTAZZI M E.	
PA	PA	(MEND/) MENDEZ S.	
XX	PI	Hotez P, Ashcom J, Edamchian M, Zhan B, Wang Y, Hawdon J;	
PI	PI	Loukas A, Williamson A, Jones B, Bethony J, Goud G, Bottazzi ME;	
PI	PI	Mendez S;	
XX	DR	WPI; 2005-194935/20.	
DR	DR	N-PSDB; ADY37229.	
XX	PT	Composition useful for vaccinating or eliciting immune response against	
PT	PT	hookworm in mammal, comprises copy of recombinant or synthetic antigen or	
PT	PT	their fragments derived from hookworm, and carrier.	
XX	PS	Disclosure; SEQ ID NO 2; 227pp; English.	
XX	CC	The invention relates to a composition (I) comprising a copy of	
CC	CC	recombinant or synthetic antigen or their fragments derived from hookworm	
CC	CC	and a carrier. (I) is useful for vaccinating or eliciting an immune	







SQ Sequence 425 AA;  
 Query Match 16.8%; Score 211; DB 5; Length 425;  
 Best Local Similarity 26.5%; Pred. No. 2.2e-12;  
 Matches 65; Conservative 36; Mismatches 76; Indels 68; Gaps 10;  
 QY 3 AAVVAVLLALFSYAEA-GFCCPNSLSQSDSARQIFLDFHNDVRNIALG-----NGLIN 56  
 DB 2 AVLAVVLLACLIERAVAQTFGCSNT-KINDQARKMFYDAHNDARRSMAKGLPNKCGL-- 58  
 QY 57 WTVNADAVILGPAQNMVKVDWDCNLEBVAQCIAPC-----ND 94  
 DB 59 -----LSGKKNVYELNWDCEMEAKAQEWADGCPSSFQTFDPTWGQNYATYMGSIAD 109  
 QY 95 PLPINTSLAQNIALRWLYFKDSEETVLQOVSWYVVSASLGFPMKGTKLDQFANQWAEPLAN 154  
 DB 110 PLPY-ASMAVN-----GWSSEIRTVGLTDPD--NKYTNSAMFRFAN 147  
 QY 155 IANYRNRYGCAHKICPAQQNMVSCVYGSFKLAPNEVIMQEGKACVCDAR---PDSFC 210  
 DB 148 MANGKASAFGCAVALCAGK--LSINCIYKIGYMTNATIEYKGDACTSDAECTTYSDSQC 205  
 QY 211 CDNLC 215  
 DB 206 KNGLC 210  
 RESULT 10  
 ID ADY37254 standard; protein; 424 AA.  
 AC ADY37254;  
 XX 05-MAY-2005 (first entry)  
 DE Hookworm antigen #13.  
 XX  
 KW antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
 KW virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
 KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
 KW tetanus; diphtheria; pertussis; polio; gene; ds.  
 OS Ancylostoma caninum.  
 PN US2005042232-A1.  
 XX  
 PD 24-FEB-2005.  
 XX  
 PF 16-APR-2004; 2004US-00825692.  
 XX  
 PR 17-OCT-2001; 2001US-0329533P.  
 PR 23-NOV-2001; 2001US-0332007P.  
 PR 26-APR-2002; 2002US-0375404P.  
 PR 17-OCT-2002; 2002WO-US033106.  
 PR 26-SEP-2003; 2003US-0505848P.  
 XX  
 PA (HOTE/) HOTEZ P.  
 PA (ASHC/) ASHCOM J.  
 PA (BDAM/) BDAMCHIAN M.  
 PA (ZHAN/) ZHAN B.  
 PA (WANG/) WANG Y.  
 PA (HAWD/) HAWDON J.  
 PA (LOUK/) LOUKAS A.  
 PA (WILL/) WILLIAMSON A.  
 PA (JONE/) JONES B.  
 PA (BETH/) BETHONY J.  
 PA (GOUT/) GOUT G.  
 PA (BOT/) BOTTAZZI M E.  
 PA (MEND/) MENDEZ S.  
 XX  
 PI Hotez P, Ashcom J, Bdamchian M, Zhan B, Wang Y, Hawdon J,  
 PI Loukas A, Williamson A, Jones B, Bethony J, Goud G, Bottazzi ME,  
 PI Mendez S;

XX WPI; 2005-194935/20.  
 DR N-PSDB; ADY37253.  
 XX  
 PT Composition useful for vaccinating or eliciting immune response against  
 PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or  
 PT their fragments derived from hookworm, and carrier.  
 XX  
 PS Disclosure; SEQ ID NO 26; 227pp; English.  
 XX  
 CC The invention relates to a composition (I) comprising a copy of  
 CC recombinant or synthetic antigen or their fragments derived from hookworm  
 CC and a carrier. (I) is useful for vaccinating or eliciting an immune  
 CC response against hookworm in a mammal, reducing blood loss in a patient  
 CC infected with hookworm, reducing hookworm size or quantitative egg count  
 CC or hookworm burden in a patient infected with hookworm, or decreasing L3  
 CC migration across skin of a mammal. The above method further involves  
 CC chemically treating a hookworm-infected patient prior to the step of  
 CC administering. (M1) is also useful for vaccinating patient against  
 CC infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,  
 CC diphtheria, pertussis or polio. This sequence corresponds to a protein of  
 CC the invention.  
 XX  
 SQ Sequence 424 AA;  
 Query Match 15.2%; Score 191; DB 9; Length 424;  
 Best Local Similarity 29.5%; Pred. No. 2.3e-10;  
 Matches 62; Conservative 31; Mismatches 83; Indels 34; Gaps 11;  
 QY 23 CPNSLSQSDSARQIFLDFHNDVRNIALGNGLINWTVNADAVILGPAQNMVKVDWDCNLE 82  
 DB 21 CPGN-DLTDARTLLTRVHNSIRREIA--QGVAN--NYHGGKLPAGKNYRMRYSCELE 74  
 QY 83 EVA--AQCIAPCNDPLPINTSLAQNIALRWLY-----FKDSEETVLQOVSWYVVSASL 133  
 DB 75 QAALDASQTF-CSASLSEEPQKYGQNIQAVTPTSIARPKDULLEDAVKQ---WYLPVITY 130  
 QY 134 GFPMKGTKLDQFANQWAEPLANIANRYNRKRYGCAHKICPAQQNMVSCVYGSFKLAP 189  
 DB 131 GQRA-----ANKFTDPRLYTPANLAYDKNTALGCHYAKCGDPDRIVISCMYNN-VVPD 183  
 QY 190 NEVIMQEGKACV---CDARPSFCCDNLC 215  
 DB 184 NAVIYEPGTACVKDADCTTYPOSTCKDSL 213  
 RESULT 11  
 ID ADY37256 standard; protein; 451 AA.  
 AC ADY37256;  
 XX 05-MAY-2005 (first entry)  
 DE Hookworm antigen #14.  
 XX  
 KW antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
 KW virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
 KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
 KW tetanus; diphtheria; pertussis; polio; gene; ds.  
 OS Ancylostoma caninum.  
 PN US2005042232-A1.  
 XX  
 PD 24-FEB-2005.  
 XX  
 PF 16-APR-2004; 2004US-00825692.  
 XX  
 PR 17-OCT-2001; 2001US-0329533P.  
 PR 23-NOV-2001; 2001US-0332007P.  
 PR 26-APR-2002; 2002US-0375404P.  
 PR 17-OCT-2002; 2002WO-US033106.



PR 26-SEP-2003; 2003US-0505848P.

XX	(HOTE/)	HOTEZ P.
PA	(ASHC/)	ASHCOM J.
PA	(BDAM/)	BDAMCHIAN M.
PA	(ZHAN/)	ZHAN B.
PA	(WANG/)	WANG Y.
PA	(HAMW/)	HAMDON J.
PA	(LOUK/)	LOUKAS A.
PA	(WILL/)	WILLIAMSON A.
PA	(JONE/)	JONES B.
PA	(BETH/)	BETHONY J.
PA	(GOUT/)	GOUT G.
PA	(BOTI/)	BOTTAZZI M E.
PA	(MEND/)	MELENZ S.

22 Ashcom J, Bdarchian M, Zhan B, Wang Y, Hawdon J;  
PI Horez P, Ashcom J, Bdarchian M, Zhan B, Wang Y, Hawdon J;  
PI Loukas A, Williamson A, Jones B, Bethony J, Goud G, Bottazzi ME;  
PI Mendez S;

DR WPI; 2005-194935/20.  
DR N-PSDB; ADY37255.

Composition useful for vaccinating or eliciting immune response against PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or their fragments derived from hookworm, and carrier.

PS Disclosure; SEQ ID NO 28; 227pp; English.

The invention relates to a composition (I) comprising a copy of recombinant or synthetic antigen or their fragments derived from hookworm and a carrier. (I) is useful for vaccinating or eliciting an immune response against hookworm in a mammal, reducing blood loss in a patient infected with hookworm, reducing hookworm size or quantitative egg count or hookworm burden in a patient infected with hookworm, or decreasing L3 migration across skin of a mammal. The above method further involves chemically treating a hookworm-infected patient prior to the step of administering. (M) is also useful for vaccinating patient against infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus, diphtheria, pertussis or polio. This sequence corresponds to a protein of the invention.

Sequence 451 AA;

Query Match	14.9%;	Score 186.5;	DB 9;	Length 451;
Best Local Similarity	27.1%;	Pred. No. 7.2e-10;		
Matches 59;	Conservative 39;	Mismatches 95;	Indels 25;	Gaps 9;

[illegible]

## RESULT 12

ADY37286  
ID ADY37286 standard; protein; 217 AA.

AC ADY37286;

DT 05-MAY-2005 (first entry)

DE Hookworm antigen #29.

XX antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
KM virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
KM hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
KM tetanus; diphtheria; pertussis; polio; gene; ds.

OS *Ancylostoma ceylanicum*.

PN US2005042232-A1.

PD 24-FEB-2005.

PF 16-APR-2004; 2004US-00825692.

PR 17-OCT-2001; 2001US-0329533P.  
PR 23-NOV-2001; 2001US-0332007P.  
PR 26-APR-2002; 2002US-0375404P.  
PR 17-OCT-2002; 2002WO-03734106.  
PR 26-SEP-2003; 2003US-0505848P.

PA	(HOTE//)	HOTEZ P.
PA	(ASHC//)	ASHCOM J.
PA	(BDAM//)	BDAMCHIAN M.
PA	(ZHAN//)	ZHAN B.
PA	(WANG//)	WANG Y.
PA	(HAMD//)	HAMDON J.
PA	(LOUK//)	LOUKAS A.
PA	(WILL//)	WILLIAMSON A.
PA	(JONE//)	JONES B.
PA	(BETH//)	BETHONY J.
PA	(GOND//)	GOND G.
PA	(BOT//)	BOTTAZZI M E
PA	(MEND//)	MELENDEZ S.

PI	Hotez P,	Ashcom J,	Bdamchian M,	Zhan B,	Wang Y,	Hawdon J;
PI	Loukas A,	Williamson A,	Jones B,	Bethony J,	Goud G,	Bottezzi ME;
PI	Mendez S;					

DR WPI; 2005-194935/20.  
DR N-PSDB; ADY37285.

Composition useful for vaccinating or eliciting immune response against hookworm in mammal, comprises copy of recombinant or synthetic antigen or their fragments derived from hookworm, and carrier.

PS Disclosure; SEQ ID NO 58; 227pp; English.

The invention relates to a composition (I) comprising a copy of recombinant or synthetic antigen or their fragments derived from hookworm and a carrier. (I) is useful for vaccinating or eliciting an immune response against hookworm in a mammal, reducing blood loss in a patient infected with hookworm, reducing hookworm size or quantitative egg count or hookworm burden in a patient infected with hookworm, or decreasing L3 migration across skin of a mammal. The above method further involves chemically treating a hookworm-infected patient prior to the step of administering. (M1) is also useful for vaccinating patient against infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus, diphtheria, pertussis or polio. This sequence corresponds to a protein of the invention.

Sequence 217 AA;

Query Match	13.2%;	Score 166;	DB 9;	Length 217;
Best Local Similarity	27.5%;	Pred. No. 3.4e-08;		
Matches	63;	Conservative	28;	Mismatches 106;
				Indels 32;
				Gaps 8;

```
QY      5 VVVAVLLALFSYAEGFCCPNLSQSDSARQIFLDFHNDVRRNIALGNGLINVTYNADAV   64
        | : ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      3 VPLIVLLAVSDANSVRCGNNGM--TDEARQKFLDMHNGYRSQVAKGQAKDALSNAAP--   58
```

QY 65 ILGPAQNMVKKVWDWCNLEEEVAAQQIAPCNDPLPINTSLAQNIARWL-YFKDSSEETVLQO 123

Db 59 ---KAAKMKKVVYDCVESTAMQNAKCVFTHSHMKLGENTI--WTTAREMDKVKSAEQ 113  
QY 124 VSWVYVSAASLGFMKGTKLDFANQWAEPLANIANRYNR-----KVGCAHKICPAQQNMV 177  
Db 114 ASQGWFESELAIEYGVGPENKLTMTQLMNRPNYTIQIHTQMTWQDTYKLGCVVEWCS----- 167  
QY 178 VSCVYGSPELAF-----NEVIWQEGKACVCDARPDSPCCDNLCDTRDAA 221  
Db 168 -SMTYGVCOYSPQGNMMSIIEKGNPCT----QSDDCGSNARCTADKA 211  
RESULT 13  
AAR26786  
ID AAR26786 standard; protein; 440 AA.  
XX  
AC AAR26786;  
XX  
DT 25-MAR-2003 (revised)  
DT 26-JAN-1993 (first entry)  
XX  
DE Sequence of a homologue of the 45 kD antigen encoded by pBTA963.  
XX  
KW Antigen; vaccine; parasitic nematode; immunogen.  
XX  
OS Haemonchus contortus.  
FH  
XX  
FH Key Location/Qualifiers  
FT Region 12..31  
FT /notes "homologous to 45 kD protein"  
FT 239..256  
FT /notes "see above"  
XX  
PN WO9213889-A1.  
XX  
PD 20-AUG-1992.  
XX  
PF 06-FEB-1992; 92WO-AU000040.  
XX  
PR 06-FEB-1991; 91AU-00004486.  
XX  
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
XX  
PI Sharp PJ, Wagland BM, Cobon GS;  
XX  
DR WPI, 1992-299990/36.  
DR N-PSDB; AAQ28094.  
XX  
PT Vaccines comprise antigens derived from parasitic nematodes - useful for  
PT passive immunisation against roundworm, whipworm, filarial worm,  
PT threadworm and hookworm.  
XX  
PS Claim 11; Fig 8; 86pp; English.  
XX  
CC PCR primers A112/301 and A112/302 were designed to be used to amplify the  
CC DNA encoding the 45 kD protein. In the reaction containing primer  
CC A112/301 a unique band of approximately 650 bp was observed. The  
CC approximately 650 bp band was not seen when primer A112/302 was used. The  
CC PCR cloned was grown and purified for use as a hybridisation probe to  
CC screen the pBTA879 cDNA library. pBTA879 is the sequence of one of these  
CC clones. There is no methionine initiation codon present in this region of  
CC the sequence so this clone probably does not represent the complete  
CC coding region. pBTA879 probably encodes a glycoprotein which is related  
CC to, but not identical to, the native glycoprotein isolated from H.  
CC contortus. In order to isolate a cDNA clone coding for the full length  
CC native 45 kD antigen, cDNA libraries were screened with the fragment  
CC isolated from pBTA879. AAQ28094 is the sequence of one of these clones,  
CC pBTA 963. Once again, this clone does not contain an initiation  
CC methionine. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-  
CC MAR-2003 to correct PA field.)  
XX  
SQ Sequence 440 AA;

Query Match 13.0%; Score 163; DB 2; Length 440;  
Best Local Similarity 26.6%; Pred. No. 1.7e-07;  
Matches 59; Conservative 24; Mismatches 109; Indels 30; Gaps 9;  
QY 15 SYAAGECCPNLSQSDSARQIFLDFHNDVRRNIALGNGLINWTVNADAVILGPAQNMVYK 74  
Db 7 SCLKAFCEPTSDNGMTDEIRQIFVDKHNERYSLIAKGQ-----AKNKLGGFAPKARMLK 61  
QY 75 VDWDCNLEVAQAQIAPC--NDPLPINTSLAQNIAEWL-----YFKDSEETVLQOVSF 126  
Db 62 VGIDCEVEANTAAAYAKECKFEHDPPEQRNYWGQNL--WMLGGTNYSK--TESALLSVQAW 117  
QY 127 YWVSASLG-----FMKGTKLDFANQWAEPLANIANRYNRKVGCAHKICPAQQNMVYSCV 181  
Db 118 YWELKMFVDPDENILTMEVFDKGVGHYTQ---VAMQSSDKIGCAVEWCPTM--TLVACB 171  
QY 182 YGSPKLAPEVWQEGKACVCDARPDSPCCDNLCDTRDAAV 223  
Db 172 YNPAGNRINHYYIDIGDPCCTD--EDWQCTGCTCSKDEALCI 211  
RESULT 14  
ADY37296  
ID ADY37296 standard; protein; 216 AA.  
XX  
AC ADY37296;  
XX  
DT 05-MAY-2005 (first entry)  
XX  
DE Hookworm antigen #34.  
XX  
KW antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
KW virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
KW tetanus; diphtheria; pertussis; polio.  
XX  
OS Ancylostoma duodenale.  
XX  
PN US2005042232-A1.  
XX  
PD 24-FEB-2005.  
XX  
PF 16-APR-2004; 2004US-00825692.  
XX  
PR 17-OCT-2001; 2001US-0329533P.  
PR 23-NOV-2001; 2001US-0332007P.  
PR 26-APR-2002; 2002US-0375404P.  
PR 17-OCT-2002; 2002WO-US033106.  
PR 26-SEP-2003; 2003US-0505848P.  
XX  
PA (HOTE/) HOTEZ P.  
PA (ASHC/) ASHCOM J.  
PA (BDAM/) BDAMCHIAN M.  
PA (ZHAN/) ZHAN B.  
PA (WANG/) WANG Y.  
PA (HAWD/) HAWDON J.  
PA (LOUK/) LOUKAS A.  
PA (WILL/) WILLIAMSON A.  
PA (JONE/) JONES B.  
PA (BETH/) BETHONY J.  
PA (GOUT/) GOUT G.  
PA (BOT/) BOTTAZZI M B.  
PA (MEN/) MENDEZ S.  
XX  
PI Hotez P, Ashcom J, Bdanchian M, Zhan B, Wang Y, Hawdon J,  
PI Loukas A, Williamson A, Jones B, Bethony J, Goud G, Bottazzi MB,  
PI Mendez S;  
XX  
DR WPI, 2005-194935/20.  
XX  
PT Composition useful for vaccinating or eliciting immune response against  
PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or  
PT their fragments derived from hookworm, and carrier.

XX Disclosure; SEQ ID NO 68; 227pp; English.  
PS  
XX  
CC The invention relates to a composition (I) comprising a copy of  
CC recombinant or synthetic antigen or their fragments derived from hookworm  
CC and a carrier. (I) is useful for vaccinating or eliciting an immune  
CC response against hookworm in a mammal, reducing blood loss in a patient  
CC infected with hookworm, reducing hookworm size or quantitative egg count  
CC or hookworm burden in a patient infected with hookworm, or decreasing L3  
CC migration across skin of a mammal. The above method further involves  
CC chemically treating a hookworm-infected patient prior to the step of  
CC administering. (M1) is also useful for vaccinating patient against  
CC infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,  
CC diphtheria, pertussis or polio. This sequence corresponds to a protein of  
CC the invention.  
XX  
SQ Sequence 216 AA;  
  
Query Match 12.8%; Score 160; DB 9; Length 216;  
Best Local Similarity 27.4%; Pred. No. 1.4e-07;  
Matches 68; Conservative 31; Mismatches 85; Indels 64; Gaps 15;  
  
QY 5 VVAVVLLALFSYAAG--FCCPNSLSQSDSARQIFLDFHNDVRRNIALGGLINWTVNAD 62  
Db 2 LVPVALLALLAVAVEGNSMRCGN--GMTDEARQEFLDVHNGYSKVAKGQA-----KD 53  
  
QY 63 AVILG---PAQNMVYKVDMDCNLEVAQAQIAPC---NDPLPINTSLAQNIARWLYFKDS 115  
Db 54 A--LGGNAPKAAKMKMYDCVVESTAMQDAKKCVFAHSHKGLGENIYMTARQM----- 106  
  
QY 116 EETVLQQVS--WYWSASLGFMKGTKLD-QFANQ-----WAEPLANIANYRNR 161  
Db 107 DKAEAAQQASDGMFAELAKYGVQENKLTMLMNRGVNIGHYTQMWQF-----SY 157  
  
QY 162 KVGCAHKICPAQOMNVSCVYGSFKLAP----NEVIWQEGKACVCDARPDSPCCDNL-C 215  
Db 158 KLGCVYEWCP-----SMTYGVQYSPQGNMMSIYEKGNPCT---QSDCGSNAKC 206  
  
QY 216 DTRDAASV 223  
Db 207 SSGEALCI 214  
  
RESULT 15  
ADY37297  
ID ADY37297 standard; protein; 207 AA.  
XX  
AC ADY37297;  
XX  
DT 05-MAY-2005 (first entry)  
XX  
DE Hookworm antigen #35.  
XX  
KW antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
KW virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
KW tetanus; diphtheria; pertussis; polio.  
XX  
OS Necator americanus.  
XX  
PN US2005042232-A1.  
XX  
PD 24-FEB-2005.  
XX  
PF 16-APR-2004; 2004US-00825692.  
XX  
PR 17-OCT-2001; 2001US-0329533P.  
PR 23-NOV-2001; 2001US-0332007P.  
PR 26-APR-2002; 2002US-0375404P.  
PR 17-OCT-2002; 2002WO-US033106.  
PR 26-SEP-2003; 2003US-0505848P.  
XX  
PA (HOTE/) HOTEZ P.

PA (ASHC/) ASHCOM J.  
PA (BDAM/) BDAMCHIAN M.  
PA (ZHAN/) ZHAN B.  
PA (WANG/) WANG Y.  
PA (HAWD/) HAWDON J.  
PA (LOUK/) LOUKAS A.  
PA (WILL/) WILLIAMSON A.  
PA (JONE/) JONES B.  
PA (BETH/) BETHONY J.  
PA (GOUT/) GOUT G.  
PA (BOT/) BOTTAZZI M E.  
PA (MEN/) MENDEZ S.  
XX  
PI Hotez P, Ashcom J, Bdamchian M, Zhan B, Wang Y, Hawdon J;  
PI Loukas A, Williamson A, Jones B, Bethony J, Goud G, Bottazzi ME;  
PI Mendez S;  
XX  
DR WPI; 2005-194935/20.  
XX  
PT Composition useful for vaccinating or eliciting immune response against  
PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or  
PT their fragments derived from hookworm, and carrier.  
XX  
XX Disclosure; SEQ ID NO 69; 227pp; English.  
XX  
PS The invention relates to a composition (I) comprising a copy of  
XX recombinant or synthetic antigen or their fragments derived from hookworm  
CC and a carrier. (I) is useful for vaccinating or eliciting an immune  
CC response against hookworm in a mammal, reducing blood loss in a patient  
CC infected with hookworm, reducing hookworm size or quantitative egg count  
CC or hookworm burden in a patient infected with hookworm, or decreasing L3  
CC migration across skin of a mammal. The above method further involves  
CC chemically treating a hookworm-infected patient prior to the step of  
CC administering. (M1) is also useful for vaccinating patient against  
CC infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,  
CC diphtheria, pertussis or polio. This sequence corresponds to a protein of  
CC the invention.  
XX  
SQ Sequence 207 AA;  
  
Query Match 12.3%; Score 154; DB 9; Length 207;  
Best Local Similarity 26.8%; Pred. No. 5.3e-07;  
Matches 62; Conservative 35; Mismatches 84; Indels 50; Gaps 14;  
  
QY 3 AAVVAVVLLALFSYAAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGGLINWTVNAD 62  
Db 2 SSITCLVLLSLIAYSKAG--CPDN-GMSEARQKFLHNSLRSSVALGQAKDGAGGNAP 58  
  
QY 63 AVILGPAQNMVYKVDMDCNLEVAQAQIAPC---NDPLPINTSLAQNIARWLYFKDSHEE 118  
Db 59 -----KAAKMTMAYDCEVEKTMNNNAKQCVFKHSQPNQ-RKGLGENI---FMSSDSGKA 109  
  
QY 119 TVLQQVS--WYWSASLGFMKGTKL--DOFAN-----QWAEPLANIANYRNRKVGCA 166  
Db 110 KAAEQASKAWFGELEKGVGNLKLTTGLFSRGVGHYTQMWQETV-----KLGCY 160  
  
QY 167 HKICPAQOMN-VVSCVYGSFKLAPNEVIWQEGKACVCDARPDSPCCDNLCD 216  
Db 161 VEAC---SNMCTVVCQYGPAGNMGMKDIYKGEPC-----SKCEN-CD 199

Search completed: March 3, 2006, 19:05:33  
Job time : 191 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:09:42 ; Search time 47 Seconds  
(without alignments)  
406.342 Million cell updates/sec

Title: US-10-527-771-10

Perfect score: 1253  
Sequence: 1 MSAAVVAVLALFSYAAG.....DNLCDTRDASVRHQCASP 231

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/6\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/H\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/RE\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	22.1	424	1	US-08-419-414-2 Sequence 2, Appli
2	165	13.2	440	1	US-07-930-686-12 Sequence 12, Appl
3	165	13.2	440	1	US-08-460-998-12 Sequence 12, Appl
4	137	10.9	227	2	US-08-450-482B-125 Sequence 125, App
5	137	10.9	227	2	US-08-151-064D-123 Sequence 123, App
6	137	10.9	229	2	US-08-450-482B-127 Sequence 127, App
7	137	10.9	229	2	US-08-151-064D-125 Sequence 125, App
8	137	10.9	238	2	US-08-450-482B-117 Sequence 117, App
9	137	10.9	238	2	US-08-151-064D-115 Sequence 115, App
10	136.5	10.9	248	2	US-08-450-482B-115 Sequence 115, App
11	136.5	10.9	248	2	US-08-151-064D-113 Sequence 113, App
12	132.5	10.6	248	4	PCT-US96-07709-30 Sequence 30, Appl
13	132.5	10.6	248	4	PCT-US96-07709-25 Sequence 25, Appl
14	131	10.5	454	1	US-07-930-686-10 Sequence 10, Appl
15	131	10.5	454	1	US-08-460-998-10 Sequence 10, Appl
16	130.5	10.4	223	2	US-09-541-759-1 Sequence 1, Appli
17	130	10.4	232	2	US-08-450-482B-123 Sequence 123, App
18	130	10.4	232	2	US-08-151-064D-121 Sequence 121, App
19	128.5	10.3	253	2	US-09-541-759-2 Sequence 2, Appli
20	126	10.1	270	2	US-08-450-482B-114 Sequence 114, App
21	126	10.1	270	2	US-08-151-064D-112 Sequence 112, App
22	119	9.5	455	2	US-09-866-028-50 Sequence 50, Appl
23	119	9.5	455	2	US-09-944-457-50 Sequence 50, Appl
24	119	9.5	455	2	US-09-945-584-50 Sequence 50, Appl
25	119	9.5	455	2	US-09-944-944-50 Sequence 50, Appl
26	119	9.5	455	2	US-09-945-587-50 Sequence 50, Appl
27	118	9.4	204	4	PCT-US96-07709-33 Sequence 33, Appl

28	117.5	9.4	220	2	US-09-541-759-3 Sequence 3, Appli
29	117	9.3	224	1	US-08-173-510B-87 Sequence 87, Appl
30	117	9.3	224	1	US-08-458-218-85 Sequence 85, Appl
31	117	9.3	224	1	US-08-450-497-87 Sequence 87, Appl
32	117	9.3	224	2	US-08-060-433C-37 Sequence 37, Appl
33	117	9.3	224	2	US-08-450-482B-87 Sequence 87, Appl
34	117	9.3	224	2	US-08-151-064D-85 Sequence 85, Appl
35	115.5	9.2	192	2	US-09-541-759-5 Sequence 5, Appli
36	115.5	9.2	231	1	US-08-173-510B-86 Sequence 86, Appl
37	115.5	9.2	231	1	US-08-458-218-84 Sequence 84, Appl
38	115.5	9.2	231	1	US-08-450-497-86 Sequence 86, Appl
39	115.5	9.2	231	2	US-08-060-433C-36 Sequence 36, Appl
40	115.5	9.2	231	2	US-08-450-482B-86 Sequence 86, Appl
41	115.5	9.2	231	2	US-08-151-064D-84 Sequence 84, Appl
42	113	9.0	274	1	US-08-173-510B-83 Sequence 83, Appl
43	113	9.0	274	1	US-08-458-218-81 Sequence 81, Appl
44	113	9.0	274	1	US-08-450-497-83 Sequence 83, Appl
45	113	9.0	274	2	US-08-060-433C-33 Sequence 33, Appl

## ALIGNMENTS

RESULT 1  
US-08-419-414-2  
Sequence 2, Application US/08419414  
Patent No. 5753787  
GENERAL INFORMATION:  
APPLICANT: Hawdon, John M.  
APPLICANT: Hotez, Peter J.  
APPLICANT: Jones, Brian F.  
TITLE OF INVENTION: Hookworm Vaccine  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/419,414  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: YU113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 873-8795  
TELEFAX: (404) 873-8795  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULAR TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Ancylostoma caninum  
DEVELOPMENTAL STAGE: Larva  
US-08-419-414-2  
Query Match 22.1%; Score 277; DB 1; Length 424;  
Best Local Similarity 30.9%; Pred. No. 2.2e-21;

Matches	76;	Conservative	35;	Mismatches	97;	Indels	38;	Gaps	11;
QY	1	MSAAVVAVLLAL----	FSYAEAGFCCPNLSQSDSARQIFLDFHNDVRRNIALG	IN	56				
Db	1	MSPVIVSVIFTIAFC	DASPARDFGCSNS-GITDKDQAF	LDHFNNARRVAKG----	54				
QY	57	WTVNADAVILGPAQNM	YKVDMDCNLEEVAAQIAPCNDPLPINTSLAQN	IARWLY---FK	113				
Db	55	-VEDSNGKLNPAKNN	MYKLSNDCAMEQQLQDAIQSCPSAFAGIOGVAQN	VMSSSSG	FP	113			
QY	114	DSEETVLQQVSWY	WVSASLGFNKGTKLDQ-----FANQWAEPLAN	IANYRRNKV	GCA	166			
Db	114	DPSVK-IEQTL	SGWWSGAK--KNGVGPDNKKYNGGGLFA-----FSNMVYSE	TKLGCA	163				
QY	167	HKICPAQNMV	SVCGSPKLAPEVIMQEGKACV----CDARPD	SFCCDNLCDT-RDA	221				
Db	164	YKVCGTK--LAV	SCIYNGVGYITNQPMWETGQACKTGADCS	TYKNSGCEDGLCTKGP	DVP	221			
QY	222	SVRHQC	227						
Db	222	ETNQOC	227						

```

RESULT 2
US-07-930-686-12
; Sequence 12, Application US/07930686
; Patent No. 5525508
; GENERAL INFORMATION:
; APPLICANT: Sharp, Phillip J
; APPLICANT: Wagland, Barry M
; APPLICANT: Cobon, Gary S
; TITLE OF INVENTION: Nematode Vaccine
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley and Lardner
; STREET: suite 500, 1800 Diagonal Road
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 2213-0299
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/930,686
; FILING DATE: 19921006
; CLASSIFICATION: 514
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK4486
; FILING DATE: 06-FEB-1991
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00040
; FILING DATE: 06-FEB-1992
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A
; REGISTRATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
;
; TELEX: 899149
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; US-07-930-686-12

```

QY	15	SYAEGFCCPNLSQSDSARQIFLDFHNDVRRNIALGNGLINWTVNADAVILGPAQNMYK	74
		:     :   :       :     :	
		:     :   :       :     :	
Db	7	SLKAAFCPTSDNGMTDEIRQIFVDKNEYRSIIAKGQ----	AKNKLGGFAPKARMUK 61
QY	75	VDWDCNLBEVAAQQAIPC--NDPLPINTSLAQNIRWL-----YFKDSEETVLQQVSW	126
		:     :     :     :     :     :	
Db	62	VGYDCEVEANTAAVAYAKECKFEHDPBQNRWGNL--WMLGTTYSK--TESAKLSVQAW	117
QY	127	YWVSASLG-----FMKGTKLDQFANQWAEPLANIANYRNRKVGCAHKICPAQNMVVS	181
		:   :   :   :   :   :   :   :   :	
Db	118	YMWELKMFVDPDENILTMEVFDRGVGHYTQ----VAMQSSDKIGCAVEMCPTM--TLVACE	171
QY	182	YGSFPLAPNEVIMQEGKACVCDARPDSPCCDNLCDTRDAASV	223
		:   :   :   :   :   :   :   :   :	
Db	172	YNPAGNRINHYYIDIGDPCTTD--EDCQCTGCTCSKDEALCI	211

```

RESULT 3
US-08-460-998-12
; Sequence 12, Application US/08460998
; Patent No. 5942413
; GENERAL INFORMATION:
; APPLICANT: Sharp, Phillip J
; APPLICANT: Wagland, Barry M
; APPLICANT: Cobon, Gary S
; TITLE OF INVENTION: Nematode Vaccine
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley and Lardner
; STREET: suite 500, 3000 K Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,998
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,686
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK4486
; FILING DATE: 06-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00040
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A
; REGISTRATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-998-12

Query Match          13.2%; Score 165; DB 1; Length 440;
Best Local Similarity 26.6%; Pred. No. 3e-09;
Matches 59; Conservative 24; Mismatches 109; Indels 30; Gaps 9;

15 SYAAGFCCNSLSQSDSARQIFLDFHNDVRRNIALGNGLINWTVNADAVILGPAQNNYK 74
| : | | | : | | | | | : | | | | | | | | | | | | | | | | | | | | | |

```

```

Db      /  SCLKAAPCPTSDNGMTDEIRQIFVDDKHNBYSRIIAKQ-----AKNKGGAAPKAARMLK 61
QY      75  VDMDCNLEBEVAAQOIAPC---NDPLPINTSLAONIARWL-----YFKDSEETVLQOVSW 126
         |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      62  VGYDCEVEANTAAAYAKECKFEHDPPEQRNYWGONL--WMLGTTYSK--TESAKLSVOQAW 117
QY      127 YWVASASLG----FMKGTKLDOFANQWABEPLANIANYRNKRVGCAHKICPAQONNVVSCV 181
         ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      118 YWELKMFVGPDENILTMVEVFDRGVGHYTO----VAWQSSDKIGCAVEWCPTM--TLVACE 171
QY      182 YGSPKLAIPNEVIMQEGKACVCDARPDSPCCDNLCDTRDAASV 223
         |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      172 YNPAGNRINHHIYIDIGDPCTTD--EDCQCTGCTCSKDEALCI 211

```

[illegible]

```

Db      5 CKÖNNGSGMTNELRRRFLRLHNGYRSILALGHVNI SEESNETFLYAHRASRMKRLIDYDCDA 64

QY      82 EEVAQO-IAPCNDPLPINTSLAONIAKWL YPKDS---EEETVLQOVSWYVWSASLGFM 136
      65 EGSAYESAIKQCCSSNKSSSAEYDENV---YVIDNTYEDEVDPAKALS-SWTSQAFNLT 119

Db      137 KGTKLDPANQWAEPLANIANY---RNKRVGCAHKICPAQQNMVVS CVYGS PKLAPNEVI 193
      120 HAE--EGIPYQMNDSVSPFANVAMDAAREKLGCAVVTCDQGNTHVVCYHGPKANKEPI 177

QY      194 WQEGKAC-----VCDARPDSPCCDNLCD 216
      178 YKVGVPSCNCTEYTRGDEBEKVFCHA-DEGVCVINLRD 213

Db

```

[illegible]



Db 65 EGSAYESA IKQCSSNKSSSAEYDENV----YVIDNTYEDEVDPALKAIS-SWTSQAFNLT 119

Qy 137 KGTCLDQFANQWAEPLNIANY---RNRKVGCAHKICPAQQNMVSVSYGSPKLAENEVI 193

Db 120 HAE--EGIPYQWNDVSDFANVAMDAREKLGCAVVTCDQGNTHVCHYGPKAANKTEPI 177

Qy 194 WQEGKAC-----VCDARPDSFCCDNLCD 216

Db 178 YKVGVPSCNCTEYTRGDDEKVFCHA-DEGVGVINLRD 213

```

RESULT 6
US-08-450-482B-127
; Sequence 127, Application US/08450482B
; Patent No. 6818616
; GENERAL INFORMATION:
; APPLICANT: MATTHEW MOYLE ET AL.
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
; NUMBER OF SEQUENCES: 132
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,482B
; FILING DATE: 26-MAY-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/173,510
; FILING DATE: 23-DEC-1993
; APPLICATION NUMBER: 08/151,064
; FILING DATE: 10-NOV-1993
; APPLICATION NUMBER: 08/060,433
; FILING DATE: 11-MAY-1993
; APPLICATION NUMBER: 07/996,972
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 07/881,721
; FILING DATE: 11-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/289
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 127:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
;
US-08-450-482B-127

```

	Query Match	10.9%;	Score 137;	DB 2;	Length 229;	
	Best Local Similarity	24.9%;	Pred. No. 1.3e-06;			
	Matches	54;	Conservative	30;	Mismatches 103;	Indels 30; Gaps 8;
OY	CCPNSLSQSDSARQIFLDFHNDVRNRNIALGNGLINWTVNADAVILGPAQNMYKVDWDCL	81				
	:   ::         :     :   :   :     :					
Dd	CKQNNGSMTNELRRRFLRLANGYSITALGHVNISESNETFLYAHRASRRIRLDYDCDA	66				
OY	EEVAAGQ-IAPCNDPLPIINTSLAQNIARWLFEKDS---EEETVLQQVSWYMWVSASIGFM	136				

Db 67 EGSAYESAIIKQCCSSNNKSSSAEYDENY-----YVIDNTYEDEVDPALKAIS-SWTSQAFNLT 121

QY 137 KGTGLDQFANQWAEPLANIANY---RNRYGCAHKICPAQQNMVSVYVGSPKLAENEVI 193

Db 122 HAE--EGIPYQWNDVSDFANVAMWABAREKLGCAVTCQGNTHVCHYGPKANRTEPI 179

QY 194 WQEGKAC-----VCDARPDSPCCDNLCD 216

Db 180 YKVGVPSCSNCTEYTRGDDEKVFCHA-DEGVCVINLRD 215

RESULT 7  
US-08-151-064D-125  
Sequence 125, Application US/08151064D  
Patent No. 6962795  
GENERAL INFORMATION:  
APPLICANT: MATTHEW MOYLE ET AL.  
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITOR  
NUMBER OF SEQUENCES: 130  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/151, 064D  
FILING DATE: 10-NOVEMBER-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/060,433  
FILING DATE: 11-MAY-1993  
APPLICATION NUMBER: 07/881,721  
FILING DATE: 11-MAY-1992  
APPLICATION NUMBER: 07/996,972  
FILING DATE: 24-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 203/226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 125:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 AMINO ACIDS  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE  
US-08-151-064D-125

	Query Match	10.9%;	Score 137;	DB 2;	Length 229;	
	Best Local Similarity	24.9%;	Pred. No. 1.3e-06;			
	Matches	54;	Conservative	30;	Mismatches 103;	Indels 30; Gaps 8;
QY	22 CCPNSLSQSBSARQIFLDFHNDVRNRNALGNGLINTVTNADAVILGPAQNMYKVDWDCNL	81				
Dh	7 CKÖNNGSMTNELRRRFLRLHNGYSRIALGHVNISEESNETFYLAAHRASRMRLDYDCCA	66				
QY	82 EEVAAQQ-IAPCNDPLPINTSLAQNIAFWLYFKDS---EEETVLQGVSWYVVSASLGFWM	136				
Dh	67 EGSAYESAIKÖCCSSNKSSSAAEYDENV---YVIDNTYEDEVDPALKAIS-SWTSQAFLNLT	121				

OY 137 KGTLDQFANQWAEPLANIANY---RNRKVGCAHKICPAQONNVSCVYGS PKLAPNEVI 193  
Db 122 HAE--EGIPYQWNDVSDFANVAMDAREKLGCAVVTCDQGNTHVCHYGPKAANKTEPI 179  
OY 194 WQEGKAC-----VCDARPDSCCDNLCD 216  
Db 180 YKVGVPSCNCTEYTRGDEKEKVFCHA-DEGVCVINLRD 215

RESULT 8

US-08-450-482B-117  
; Sequence 117, Application US/08450482B  
; Patent No. 6818616  
; GENERAL INFORMATION:  
; APPLICANT: MATTHEW MOYLE ET AL.  
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
; NUMBER OF SEQUENCES: 132  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,482B  
; FILING DATE: 26-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/173,510  
; FILING DATE: 23-DEC-1993  
; APPLICATION NUMBER: 08/151,064  
; FILING DATE: 10-NOV-1993  
; APPLICATION NUMBER: 08/060,433  
; FILING DATE: 11-MAY-1993  
; APPLICATION NUMBER: 07/996,972  
; FILING DATE: 24-DEC-1992  
; APPLICATION NUMBER: 07/881,721  
; FILING DATE: 11-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 213/289  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 117:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 238 AMINO ACIDS  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PEPTIDE  
; US-08-450-482B-117

Query Match 10.9%; Score 137; DB 2; Length 238;  
Best Local Similarity 24.9%; Pred. No. 1.4e-06;  
Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps 8;

OY 22 CCPNSLSQSDSARQIFLDHNDVRRNIALGNLINTVNADAVILGPAQNMKYKVDWDCNL 81  
Db 16 CKQNGSMTNELRRRFLRLHNGYSILALGHVNISESNETFLYAHRASRMRLDIDYDCA 75  
OY 82 EEVAAQO-IAPCNDPLPINTSLAQNIAIARWLKFKDS---EEETVLQOVSWYVWSASLGFM 136  
Db 76 EGSAYESAIIKQCSSNKSSEAYDENV---YVIDNTYEDEVDPALKAIS-SWTSQAENLT 130

OY 137 KGTLDQFANQWAEPLANIANY---RNRKVGCAHKICPAQONNVSCVYGS PKLAPNEVI 193  
Db 131 HAE--EGIPYQWNDVSDFANVAMDAREKLGCAVVTCDQGNTHVCHYGPKAANKTEPI 188  
OY 194 WQEGKAC-----VCDARPDSCCDNLCD 216  
Db 189 YKVGVPSCNCTEYTRGDEKEKVFCHA-DEGVCVINLRD 224

RESULT 9

US-08-151-064D-115  
; Sequence 115, Application US/08151064D  
; Patent No. 6962795  
; GENERAL INFORMATION:  
; APPLICANT: MATTHEW MOYLE ET AL.  
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
; NUMBER OF SEQUENCES: 130  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/151,064D  
; FILING DATE: 10-NOVEMBER-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/060,433  
; FILING DATE: 11-MAY-1993  
; APPLICATION NUMBER: 07/881,721  
; FILING DATE: 11-MAY-1992  
; APPLICATION NUMBER: 07/996,972  
; FILING DATE: 24-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIGGS, SUZANNE L.  
; REGISTRATION NUMBER: 30,158  
; REFERENCE/DOCKET NUMBER: 203/226  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 115:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 238 AMINO ACIDS  
; TYPE: AMINO ACID  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PEPTIDE  
; US-08-151-064D-115

Query Match 10.9%; Score 137; DB 2; Length 238;  
Best Local Similarity 24.9%; Pred. No. 1.4e-06;  
Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps 8;

OY 22 CCPNSLSQSDSARQIFLDHNDVRRNIALGNLINTVNADAVILGPAQNMKYKVDWDCNL 81  
Db 16 CKQNGSMTNELRRRFLRLHNGYSILALGHVNISESNETFLYAHRASRMRLDIDYDCA 75  
OY 82 EEVAAQO-IAPCNDPLPINTSLAQNIAIARWLKFKDS---EEETVLQOVSWYVWSASLGFM 136  
Db 76 EGSAYESAIIKQCSSNKSSEAYDENV---YVIDNTYEDEVDPALKAIS-SWTSQAENLT 130  
OY 137 KGTLDQFANQWAEPLANIANY---RNRKVGCAHKICPAQONNVSCVYGS PKLAPNEVI 193  
Db 131 HAE--EGIPYQWNDVSDFANVAMDAREKLGCAVVTCDQGNTHVCHYGPKAANKTEPI 188





Db 120 DPRDINHALKALISWATEAFNL-----NKTGEGVYRSILNISFANLAWDT 167

Qy 161 -RKVGCAHKICPAQNMVSCVYGSPLAPNE--VIMQEGKAC-VCDAKPPDSFCC---DN 213

Db 168 REKVGCAVVKCPSGNTHV-CHY--PKIVKEGKPIYSIGKPCRCGNDYASKFCHADEG 224

Qy 214 LC 215

Db 225 VC 226

## RESULT 12

PCT-US96-07709-30

; Sequence 30, Application PC/TUS9607709

; GENERAL INFORMATION:

; APPLICANT: Tripp, Cynthia A.

; APPLICANT: Wisniewski, Nancy

; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN

; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheridan Ross & McIntosh

; STREET: 1700 Lincoln St., Suite 3500

; CITY: Denver

; STATE: Colorado

; COUNTRY: U.S.A.

; ZIP: 80203

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/07709

; FILING DATE: 23-MAY-1996

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Connell, Gary J.

; REGISTRATION NUMBER: 32,020

; REFERENCE/DOCKET NUMBER: 2618-30-PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303) 863-9700

; TELEFAX: (303) 863-0223

; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 220 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US96-07709-30

Query Match 10.6%; Score 132.5; DB 4; Length 220;

Best Local Similarity 23.5%; Pred. No. 3.8e-06;

Matches 57; Conservative 24; Mismatches 91; Indels 71; Gaps 9;

Qy 4 AVVVAVLLALFSYABAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMTV-NAD 62

Db 9 AIVVAV-----TGYNCPGSKLTALERKCI-----VGQNNKYRSDLINGKLKRN 52

Qy 63 AVILGPAQNMVSKVWDNCNLEEVAAQIAPCNDPLPINTSLAQNIARWLYFKDS---EET 119

Db 53 GTYMPRGKNMELRWDCLE-----SSAQRWANQCFGHSPROQREG 94

Qy 120 VLOQVSWYVWSASLGFMKGTKLDOFANQWAEPLANIANYRNR----- 161

Db 95 VGENVYAYWSSVSVEGLKKTAGTDAGKSWMSBLPKL--YENNPNNMTWKVAGQGVLFHT 152

Qy 162 -----KVGAHKI-CPAQNMVSCVYGSPLAPNEVIWQEGKACVCDARPSFCCD 212

Db 153 QMAWGKTYKIGCGVATGCGDGRLLIVICHYSPGNNVGEVIYQRGNPC-----KVDKDCYT 208

Qy 213 NUC 215

Db 209 KKC 211

## RESULT 13

PCT-US96-07709-25

; Sequence 25, Application PC/TUS9607709

; GENERAL INFORMATION:

; APPLICANT: Tripp, Cynthia A.

; APPLICANT: Wisniewski, Nancy

; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN

; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheridan Ross & McIntosh

; STREET: 1700 Lincoln St., Suite 3500

; CITY: Denver

; STATE: Colorado

; COUNTRY: U.S.A.

; ZIP: 80203

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/07709

; FILING DATE: 23-MAY-1996

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Connell, Gary J.

; REGISTRATION NUMBER: 32,020

; REFERENCE/DOCKET NUMBER: 2618-30-PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303) 863-9700

; TELEFAX: (303) 863-0223

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 248 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US96-07709-25

Query Match 10.6%; Score 132.5; DB 4; Length 248;

Best Local Similarity 23.5%; Pred. No. 4.5e-06;

Matches 57; Conservative 24; Mismatches 91; Indels 71; Gaps 9;

Qy 4 AVVVAVLLALFSYABAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMTV-NAD 62

Db 37 AIVVAV-----TGYNCPGSKLTALERKCI-----VGQNNKYRSDLINGKLKRN 80

Qy 63 AVILGPAQNMVSKVWDNCNLEEVAAQIAPCNDPLPINTSLAQNIARWLYFKDS---EET 119

Db 81 GTYMPRGKNMELRWDCLE-----SSAQRWANQCFGHSPROQREG 122

Qy 120 VLOQVSWYVWSASLGFMKGTKLDOFANQWAEPLANIANYRNR----- 161

Db 123 VGENVYAYWSSVSVEGLKKTAGTDAGKSWMSBLPKL--YENNPNNMTWKVAGQGVLFHT 180

Qy 162 -----KVGAHKI-CPAQNMVSCVYGSPLAPNEVIWQEGKACVCDARPSFCCD 212

Db 181 QMAWGKTYKIGCGVATGCGDGRLLIVICHYSPGNNVGEVIYQRGNPC-----KVDKDCYT 236

Qy 213 NUC 215

Db 237 KKC 239

RESULT 14

US-07-930-686-10

; Sequence 10, Application US/07930686

; Patent No. 5525508

; GENERAL INFORMATION:



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:10:28 ; Search time 165 Seconds  
(without alignments)  
584.961 Million cell updates/sec

Title: US-10-527-771-10  
Perfect score: 1253  
Sequence: 1 MSAVVAVLALFSYAEAG.....DNLCITRDASVRHQCASP 231

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA\_Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	1253	100.0	231	4	US-10-243-319C-10 Sequence 10, Appli
2	636.5	50.8	236	4	US-10-243-319C-8 Sequence 8, Appli
3	277	22.1	424	5	US-10-825-692-18 Sequence 18, Appli
4	274	21.9	425	5	US-10-825-692-67 Sequence 67, Appli
5	272.5	21.7	218	4	US-10-051-644B-5 Sequence 5, Appli
6	263	21.0	424	5	US-10-825-692-2 Sequence 2, Appli
7	262.5	20.9	422	5	US-10-825-692-56 Sequence 56, Appli
8	216.5	17.3	473	4	US-10-051-644B-3 Sequence 3, Appli
9	211	16.8	425	4	US-10-051-644B-1 Sequence 1, Appli
10	191	15.2	424	5	US-10-825-692-26 Sequence 26, Appli
11	186.5	14.9	451	5	US-10-825-692-28 Sequence 28, Appli
12	166	13.2	217	5	US-10-825-692-58 Sequence 58, Appli
13	160	12.8	216	5	US-10-825-692-68 Sequence 68, Appli
14	154	12.3	207	5	US-10-825-692-69 Sequence 69, Appli
15	147.5	11.8	218	5	US-10-825-692-20 Sequence 20, Appli
16	135	10.8	200	5	US-10-825-692-22 Sequence 22, Appli
17	124.5	9.9	395	5	US-10-450-763-57962 Sequence 57962, A
18	123	9.8	247	6	US-11-097-143-17823 Sequence 17823, A
19	123	9.8	415	4	US-10-309-290-84 Sequence 84, Appli
20	122	9.7	245	4	US-10-051-644B-6 Sequence 6, Appli
21	122	9.7	245	3	US-09-800-198-92 Sequence 92, Appli
22	122	9.7	255	3	US-09-876-225-2 Sequence 2, Appli
23	122	9.7	255	4	US-10-417-727-2 Sequence 2, Appli
24	122	9.7	255	5	US-10-909-864A-2 Sequence 2, Appli
25	121	9.7	207	4	US-10-051-644B-7 Sequence 7, Appli
26	121	9.7	308	6	US-11-097-143-19971 Sequence 19971, A
27	121	9.7	446	3	US-09-726-643-47 Sequence 47, Appli

28	121	9.7	446	3	US-09-790-264-2	Sequence 2, Appli
29	121	9.7	446	4	US-10-042-141-47	Sequence 47, Appli
30	121	9.7	446	4	US-10-269-353-2	Sequence 2, Appli
31	121	9.7	446	4	US-10-169-395-5	Sequence 5, Appli
32	121	9.7	446	5	US-10-919-272-47	Sequence 47, Appli
33	121	9.7	446	5	US-10-900-926-2	Sequence 2, Appli
34	119	9.5	455	3	US-09-866-028-50	Sequence 50, Appli
35	119	9.5	455	3	US-09-944-449-50	Sequence 50, Appli
36	119	9.5	455	3	US-09-944-457-50	Sequence 50, Appli
37	119	9.5	455	3	US-09-944-862-50	Sequence 50, Appli
38	119	9.5	455	3	US-09-945-587-50	Sequence 50, Appli
39	119	9.5	455	3	US-09-945-015-50	Sequence 50, Appli
40	119	9.5	455	3	US-09-944-396-50	Sequence 50, Appli
41	119	9.5	455	3	US-09-944-432-50	Sequence 50, Appli
42	119	9.5	455	3	US-09-943-762-50	Sequence 50, Appli
43	119	9.5	455	3	US-09-944-654-50	Sequence 50, Appli
44	119	9.5	455	3	US-09-943-851A-50	Sequence 50, Appli
45	119	9.5	455	3	US-09-944-413-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1  
US-10-243-319C-10  
Sequence 10, Application US/10243319C  
Publication No. US20040052817A1  
GENERAL INFORMATION:  
APPLICANT: Universiteit Gent  
TITLE OF INVENTION: Ostertagia vaccine  
FILE REFERENCE: Unigent Ostertagia  
CURRENT APPLICATION NUMBER: US/10/243,319C  
CURRENT FILING DATE: 2002-09-13  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 10  
LENGTH: 231  
TYPE: PRT  
ORGANISM: Ostertagia ostertagi  
US-10-243-319C-10

Query Match 100.0%; Score 1253; DB 4; Length 231;  
Best Local Similarity 100.0%; Pred. No. 5.6e-119;  
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAVVAVLALFSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRNIALGNGLINWTVN 60  
DB 1 MSAVVAVLALFSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRNIALGNGLINWTVN 60  
QY 61 ADAVITLGAQNMKYVMDNCULEEVAQAQIAPCNDPLINTSLAQNIAKWLKFSKSEETV 120  
DB 61 ADAVITLGAQNMKYVMDNCULEEVAQAQIAPCNDPLINTSLAQNIAKWLKFSKSEETV 120  
QY 121 LQOVSWYVWSASLGFPMKGTGLDQFANQWAEPLANTANYRNKVGCAHKICPAQQNMVWSC 180  
DB 121 LQOVSWYVWSASLGFPMKGTGLDQFANQWAEPLANTANYRNKVGCAHKICPAQQNMVWSC 180  
QY 181 VYGSPLAPNEVIWQEGKACVCDARPDSCDNLCTTRDASVRHQCASP 231  
DB 181 VYGSPLAPNEVIWQEGKACVCDARPDSCDNLCTTRDASVRHQCASP 231

RESULT 2  
US-10-243-319C-8  
Sequence 8, Application US/10243319C  
Publication No. US20040052817A1  
GENERAL INFORMATION:  
APPLICANT: Universiteit Gent  
TITLE OF INVENTION: Ostertagia vaccine  
FILE REFERENCE: Unigent Ostertagia  
CURRENT APPLICATION NUMBER: US/10/243,319C  
CURRENT FILING DATE: 2002-09-13  
NUMBER OF SEQ ID NOS: 14



```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Ostertagia ostertagi
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (37)..(37)
; OTHER INFORMATION: The 'Xaa' at location 37 stands for Lys, or Gln.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121)..(121)
; OTHER INFORMATION: The 'Xaa' at location 121 stands for Asn, or Thr.
US-10-243-319C-8

Query Match          50.8%; Score 636.5; DB 4; Length 236;
Best Local Similarity 52.5%; Pred. No. 3.2e-56;
Matches 125; Conservative 35; Mismatches 67; Indels 11; Gaps 4;

QY      1 MSAVVVAV--LALFSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRIALGNGLINWT 58
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      1 MQALIGIALYLVLTSTNEAGFCPADLNQTDARXIFLDFHNOVRRIAGASPLNLT 60

QY      59 VNADAV---ILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAQNIAFWLYFKD 114
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      61 --GAVQMRNVLGPAKNMYRMDWDCNLEAKAKAMIWPCCTPPLPIDTSIPQNLAQWLFQN 117

QY      115 SEERTVLQQVSWYVWSASLGFMTGKTKLDQFANQWAEPLANIANYNRNRKVGCAHKIC--PA 172
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      118 SQEXEVLQTPTWSWVTASLRNLQPDTEANIYNWQIRPLSLIANWQNLKVGCAHKVCKRPT 177

QY      173 QQNMVWSCVYGSFKLAPNEVIWQEGKACVCDARPSFCCDNLCDTRDAASVRHQCAS 230
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      178 GTNMVSCAYGGEVLQDNEVWMDKPTCMCNAYPNPFCNNLCDTIAATLRKQPKCS 235

RESULT 3
US-10-825-692-18
; Sequence 18, Application US/10825692
; Publication No. US20050042232A1
; GENERAL INFORMATION:
; APPLICANT: Hotez, Peter
; APPLICANT: Ashcom, James
; APPLICANT: Bdamchian, Mahnaz
; APPLICANT: Zhan, Bin
; APPLICANT: Wang, Yan
; APPLICANT: Hawdon, John
; APPLICANT: Loukas, Alexander
; APPLICANT: Williamson, Angela
; APPLICANT: Jones, Brian
; APPLICANT: Bethony, Jeffrey
; APPLICANT: Goud, Gaddam
; APPLICANT: Botazzi, Maria B.
; APPLICANT: Mendez, Susana
; TITLE OF INVENTION: Hookworm Vaccine
; FILE REFERENCE: 03740007aa
; CURRENT APPLICATION NUMBER: US/10/825,692
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/329,533
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 60/332,007
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: US 60/375,404
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT US02/33106
; PRIOR FILING DATE: 2002-10-17
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Ancylostoma caninum
US-10-825-692-18
```

```
Query Match          22.1%; Score 277; DB 5; Length 424;
Best Local Similarity 30.9%; Pred. No. 2.6e-19;
Matches 76; Conservative 35; Mismatches 97; Indels 38; Gaps 11;

QY      1 MSAVVVAVLAL----FSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRIALGNGLIN 56
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      1 MSPPVIVSVIFTAFCDASPARDFGCSNS-GITDKRQAFLDHFHNNARRVAKG----- 54

QY      57 WTVNADAVILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAQNIAFWLY---FK 113
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      55 -VEDSNSGKLNPAKNMYKLSWDCAMEQQLDQAIQSCPSAFAGIQVAQNVMSWSSSGGFP 113

QY      114 DSEERTVLQQVSWYVWSASLGFMTGKTKLDQ-----PANQWAEPLANIANYNRNRKVGCA 166
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      114 DPSVK-IBQTLGWSGAK--KNGVGPDKYNGGGLFA-----FSNMVYSETTKLGCA 163

QY      167 HKICPAQNMVWSCVYGSFKLAPNEVIWQEGKACV---CDARPSFCCDNLCDT-RDAA 221
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      164 YKVCGTK--LAVSCITNGVGYITNQPMWETGQACKTGADCSYKNSGCEDEGLCTKGPDPV 221

QY      222 SVRHQC 227
        |||
Db      222 ETNQOC 227

RESULT 4
US-10-825-692-67
; Sequence 67, Application US/10825692
; Publication No. US20050042232A1
; GENERAL INFORMATION:
; APPLICANT: Hotez, Peter
; APPLICANT: Ashcom, James
; APPLICANT: Bdamchian, Mahnaz
; APPLICANT: Zhan, Bin
; APPLICANT: Wang, Yan
; APPLICANT: Hawdon, John
; APPLICANT: Loukas, Alexander
; APPLICANT: Williamson, Angela
; APPLICANT: Jones, Brian
; APPLICANT: Bethony, Jeffrey
; APPLICANT: Goud, Gaddam
; APPLICANT: Botazzi, Maria B.
; APPLICANT: Mendez, Susana
; TITLE OF INVENTION: Hookworm Vaccine
; FILE REFERENCE: 03740007aa
; CURRENT APPLICATION NUMBER: US/10/825,692
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/329,533
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 60/332,007
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: US 60/375,404
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT US02/33106
; PRIOR FILING DATE: 2002-10-17
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Ancylostoma duodenale
US-10-825-692-67

Query Match          21.9%; Score 274; DB 5; Length 425;
Best Local Similarity 30.3%; Pred. No. 5.3e-19;
Matches 74; Conservative 37; Mismatches 97; Indels 36; Gaps 11;

QY      2 SAAVVVAVLAL----FSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRIALG-----N 52
        |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      3 SSVVVISVISTIAFCDASPARASFGCSNN-GITDSRQAFLDHFHNNARRVAAQGVEDNKS 61

QY      53 GLINWTVNADAVILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAQNIAFWLY- 111
```

```
Db 62 GKLN-----PAKMYKLEWDCMEQQLQDAIQSCPGSAGIQFSQNMWSMS 110
Qy 112 --FKDSEETVLQOVSWYWSASLGFMTKTLQD-QFANQWAEPLANIANRNRKVGCAHK 168
Db 111 GGFNSSEK-IESTLSGWSGAK---NNGVSDNKYTGGLVAFSNMVFSETTKIGCAYK 166
Qy 169 ICPAQNMVSVCGSPKLAPEVIMQEGKACV----CDARPDFCCDNLCDT-RDAASV 223
Db 167 VCGTK--MATSCITNGIGYITNAPMETGQACKTGADCSITYKNSGCEDSLCTKGADVPET 224
Qy 224 RHQC 227
Db 225 NQOC 228
```

## RESULT 5

```
US-10-051-644B-5
; Sequence 5, Application US/10051644B
; Publication No. US20030126625A1
; GENERAL INFORMATION:
; APPLICANT: Liu, et al.
; TITLE OF INVENTION: Screens and Assays for Agents Useful in Controlling
; TITLE OF INVENTION: Parasitic Nematodes
; FILE REFERENCE: 2002630-0012
; CURRENT APPLICATION NUMBER: US/10/051,644B
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Clustal W
; OTHER INFORMATION: Alignment of VAP-1, VAP-2, and Selected Other
; OTHER INFORMATION: Nematode VA Proteins.
US-10-051-644B-5
```

```
Query Match 21.7%; Score 272.5; DB 4; Length 218;
Best Local Similarity 31.3%; Pred. No. 3.3e-19;
Matches 73; Conservative 35; Mismatches 88; Indels 37; Gaps 10;
```

```
Qy 1 MSAVAVVAVLAL----FSYAEAGFCCPNLSQSDSARQIFLDFHNDVRRNIALGNGLIN 56
Db 1 MSPVIVSVIFTIAFCDA SPARDFGCSNS-GITDKDQAFLDHFNNARRVAKG----- 54
Qy 57 WTVNADAVILGPAQNMVKVDWDCNLEEVAQAQIAPCNDPLPINTSLAQNIARWLY---FK 113
Db 55 -VEDSNSGKLNPAKNMVKLSWDCAMEQQLQDAIQSCPSAFAGIQGVAQNMVMSWSSSGGFP 113
Qy 114 DSEETVLQOVSWYWSASLGFMTKTLQD-----FANQWAEPLANIANRNRKVCA 166
Db 114 DPSVK-IEQTLGWSGAK---KNGVGPDKNYNGGLFA-----FSNMVSETTKLGCA 163
Qy 167 HKICPAQNMVSVCGSPKLAPEVIMQEGKACV---CDARPDFCCDNLCDT-RDAASV 215
Db 164 YKVCGTK--LAVSCITNGVGYITNQPMWETGQACKTGADCSITYKNSGCEDSLCT 214
```

## RESULT 6

```
US-10-825-692-2
; Sequence 2, Application US/10825692
; Publication No. US20050042232A1
; GENERAL INFORMATION:
; APPLICANT: Hotez, Peter
; APPLICANT: Ashcom, James
; APPLICANT: Bdamchian, Mahnaz
; APPLICANT: Zhan, Bin
; APPLICANT: Wang, Yan
; APPLICANT: Hawdon, John
; APPLICANT: Loukas, Alexander
; APPLICANT: Williamson, Angela
```

```
; APPLICANT: Jones, Brian
; APPLICANT: Bethony, Jeffrey
; APPLICANT: Goud, Gaddam
; APPLICANT: Botazzi, Maria E.
; APPLICANT: Mendez, Susana
; TITLE OF INVENTION: Hookworm Vaccine
; FILE REFERENCE: 03740007aa
; CURRENT APPLICATION NUMBER: US/10/825,692
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/329,533
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 60/332,007
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: US 60/375,404
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT US02/33106
; PRIOR FILING DATE: 2002-10-17
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Necator americanus
US-10-825-692-2
```

```
Query Match 21.0%; Score 263; DB 5; Length 424;
Best Local Similarity 30.5%; Pred. No. 6.9e-18;
Matches 73; Conservative 36; Mismatches 106; Indels 24; Gaps 9;
```

```
Qy 1 MSAVAVVAVLAL----FSYAEAGFCCPNLSQSDSARQIFLDFHNDVRRNIALGNGLIN 56
Db 1 MSPVIVSVIFTIAFCNAPSARDFGCSNS-GITDSDQAFLDHFNNARRVAKG----- 54
Qy 57 WTVNADAVILGPAQNMVKVDWDCNLEEVAQAQIAPCNDPLPINTSLAQNIARWLYFKDSE 116
Db 55 -LEDNSGKLNPAKNMVKLSWDCAMEQQLQDAIQSCPSAFAGIQGVAQNMVMSWSSSGGYP 113
Qy 117 EETVLQO--VSWYWSASLGFMTKTLQD-QFANQWAEPLANIANRNRKVCAHKICPAQ 173
Db 114 DPSVKIEPTLSGWSGAK---KNGVGPDKNYTGGLFAFSNMVYSETTKLGCAKVCGTK 170
Qy 174 QNMVSVCGSPKLAPEVIMQEGKACV---CDARPDFCCDNLCDT-RDAASV RHQC 227
Db 171 --LAVSCITNGVGYITNQPMWETGQACQGTGADCSITYKNSGCEDSLCTKGPDPVETNQC 227
```

## RESULT 7

```
US-10-825-692-56
; Sequence 56, Application US/10825692
; Publication No. US20050042232A1
; GENERAL INFORMATION:
; APPLICANT: Hotez, Peter
; APPLICANT: Ashcom, James
; APPLICANT: Bdamchian, Mahnaz
; APPLICANT: Zhan, Bin
; APPLICANT: Wang, Yan
; APPLICANT: Hawdon, John
; APPLICANT: Loukas, Alexander
; APPLICANT: Williamson, Angela
; APPLICANT: Jones, Brian
; APPLICANT: Bethony, Jeffrey
; APPLICANT: Goud, Gaddam
; APPLICANT: Botazzi, Maria E.
; APPLICANT: Mendez, Susana
; TITLE OF INVENTION: Hookworm Vaccine
; FILE REFERENCE: 03740007aa
; CURRENT APPLICATION NUMBER: US/10/825,692
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/329,533
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 60/332,007
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: US 60/375,404
```







Db 59 ---KAAKMKMYVDCGVESTAMQNAKCVFTHSHMKLGENI--WMTTAREMDKYKSAEQ 113  
QY 124 VSMYVWVASLGFMKGTLDQFANQWAEPLNIANYRNR-----KVGCAHKICPAQQNMV 177  
Db 114 ASQGWSELAEGVGPENKLTMOJLNRPNTOIGHYTMVQDITYKLGCYVENCS----- 167  
QY 178 VSCVYSPKLAPE---NEVIWQEGKACVCDARPDSCCDNLCDTRDAA 221  
Db 168 -SMTYGVQCQSPQGNMNSIIEYKGNPCT----QSDCGSNARCTADKA 211

RESULT 13  
US-10-825-692-68  
; Sequence 68, Application US/10825692  
; Publication No. US20050042232A1  
; GENERAL INFORMATION:  
; APPLICANT: Hotez, Peter  
; APPLICANT: Ashcom, James  
; APPLICANT: Bdamchian, Mahnaz  
; APPLICANT: Zhan, Bin  
; APPLICANT: Wang, Yan  
; APPLICANT: Hawdon, John  
; APPLICANT: Loukas, Alexander  
; APPLICANT: Williamson, Angela  
; APPLICANT: Jones, Brian  
; APPLICANT: Bethony, Jeffrey  
; APPLICANT: Goud, Gaddam  
; APPLICANT: Botazzi, Maria E.  
; APPLICANT: Mendez, Susana  
; TITLE OF INVENTION: Hookworm Vaccine  
; FILE REFERENCE: 03740007aa  
; CURRENT APPLICATION NUMBER: US/10/825,692  
; PRIOR FILING DATE: 2004-04-16  
; PRIOR APPLICATION NUMBER: US 60/329,533  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: US 60/332,007  
; PRIOR FILING DATE: 2001-11-23  
; PRIOR APPLICATION NUMBER: US 60/375,404  
; PRIOR FILING DATE: 2002-04-26  
; PRIOR APPLICATION NUMBER: PCT US02/33106  
; PRIOR FILING DATE: 2002-10-17  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 68  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Ancylostoma duodenale  
US-10-825-692-68

Query Match 12.8%; Score 160; DB 5; Length 216;  
Best Local Similarity 27.4%; Pred. No. 9.1e-08;  
Matches 68; Conservative 31; Mismatches 85; Indels 64; Gaps 15;  
QY 5 VVAVVALLALFSYAEAG--FCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMTVNAD 62  
Db 2 LVPVALLALLAFAVEGNSMRCCGN--GMTDEARQEFLLDVHNGYRSKVAKGQA-----KD 53  
QY 63 AVILG---PAQNMVYKVDWDCNLEEVAAQOIAPC---NDPLPINTSLAQNIARWLYFKDS 115  
Db 54 A--LGNAPKAAKMKMYIDCNVESTAMQDAKCVFAHSHKLGGENIYMTARQM----- 106  
QY 116 EETVLQOVS--WYVWVASLGFMKGTLD-QFANQ-----WAEPLNIANYRNR 161  
Db 107 DKAEAAQASDGSWFABELAKYGVGOENKLTMOJLNRGVMIGHYTMVQJ-----SY 157  
QY 162 KVGCAHKICPAQQNMVSCVYSPKLAPE---NEVIWQEGKACVCDARPDSCCDNL-C 215  
Db 158 KLGCVYEWCP-----SMTYGVQCQSPQGNMNSIIEYKGNPCT----QSDCGSNAKC 206  
QY 216 DTRDASV 223  
Db 207 SSGEALCI 214

RESULT 14  
US-10-825-692-69  
; Sequence 69, Application US/10825692  
; Publication No. US20050042232A1  
; GENERAL INFORMATION:  
; APPLICANT: Hotez, Peter  
; APPLICANT: Ashcom, James  
; APPLICANT: Bdamchian, Mahnaz  
; APPLICANT: Zhan, Bin  
; APPLICANT: Wang, Yan  
; APPLICANT: Hawdon, John  
; APPLICANT: Loukas, Alexander  
; APPLICANT: Williamson, Angela  
; APPLICANT: Jones, Brian  
; APPLICANT: Bethony, Jeffrey  
; APPLICANT: Goud, Gaddam  
; APPLICANT: Botazzi, Maria E.  
; APPLICANT: Mendez, Susana  
; TITLE OF INVENTION: Hookworm Vaccine  
; FILE REFERENCE: 03740007aa  
; CURRENT APPLICATION NUMBER: US/10/825,692  
; PRIOR FILING DATE: 2004-04-16  
; PRIOR APPLICATION NUMBER: US 60/329,533  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: US 60/332,007  
; PRIOR FILING DATE: 2001-11-23  
; PRIOR APPLICATION NUMBER: US 60/375,404  
; PRIOR FILING DATE: 2002-04-26  
; PRIOR APPLICATION NUMBER: PCT US02/33106  
; PRIOR FILING DATE: 2002-10-17  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 69  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Necator americanus  
US-10-825-692-69

Query Match 12.3%; Score 154; DB 5; Length 207;  
Best Local Similarity 26.8%; Pred. No. 3.5e-07;  
Matches 62; Conservative 35; Mismatches 84; Indels 50; Gaps 14;  
QY 3 AAVVAVVALLALFSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMTVNAD 62  
Db 2 SSTICVLILSLAAYSKAG--CPDN-GMSEARQKFLHNSLRSSVALGQAKDGAGGNAP 58  
QY 63 AVILGPAQNMVYKVDWDCNLEEVAAQOIAPC---NDPLPINTSLAQNIARWLYFKDSEEE 118  
Db 59 -----KAAKMKTMAYDCEVEKTAAMNNNAKQCVFKHSQPNQ--RKGLGENI---FMSSDSGKA 109  
QY 119 TVLQOVS--WYVWVASLGFMKGTLD-QFAN-----QWAEPLNIANYRNRKVGCA 166  
Db 110 KAEQASKAWFGLAEKGVGNLKLTLGGLFSRGVGHYTMVWQJTV-----KLGCV 160  
QY 167 HKICPAQQNM-VSCVYSPKLAPENEVIWQEGKACVCDARPDSCCDNLCD 216  
Db 161 VEAC---SNMCTVVCQYGPAGNMGKDIYEKGPCT-----SKCEN-CD 199

RESULT 15  
US-10-825-692-20  
; Sequence 20, Application US/10825692  
; Publication No. US20050042232A1  
; GENERAL INFORMATION:  
; APPLICANT: Hotez, Peter  
; APPLICANT: Ashcom, James  
; APPLICANT: Bdamchian, Mahnaz  
; APPLICANT: Zhan, Bin  
; APPLICANT: Wang, Yan  
; APPLICANT: Hawdon, John  
; APPLICANT: Loukas, Alexander  
; APPLICANT: Williamson, Angela





*This Page Blank (uspto)*

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using .sw model

Run on: March 3, 2006, 19:11:22 ; Search time 21 Seconds  
(without alignments)  
220.005 Million cell updates/sec

Title: US-10-527-771-10  
Perfect score: 1253  
Sequence: 1 MSAVVAVVAVLALFSYAAG.....DNLCDTRDASVRHQCASP 231

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 135346 seqs, 20000420 residues

Total number of hits satisfying chosen parameters: 135346

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1253	100.0	231	6	US-10-527-771-10
2	636.5	50.8	236	6	US-10-527-771-2
3	130.5	10.4	223	7	US-11-153-222A-1
4	128.5	10.3	253	7	US-11-153-222A-2
5	119	9.5	243	6	US-10-515-868-5
6	117.5	9.4	220	7	US-11-153-222A-3
7	111.5	8.9	258	7	US-11-090-439-38
8	106.5	8.5	245	6	US-10-515-868-6
9	103.5	8.3	500	7	US-11-067-573-2
10	99	7.9	243	6	US-10-515-868-4
11	99	7.9	243	6	US-10-506-443A-41
12	97.5	7.8	277	6	US-10-527-500-23
13	92.5	7.4	497	6	US-10-453-372-100
14	92.5	7.4	497	6	US-10-453-372-102
15	91.5	7.3	475	6	US-10-453-372-98
16	91.5	7.3	477	6	US-10-453-372-90
17	91.5	7.3	497	6	US-10-453-372-92
18	90.5	7.2	501	6	US-10-453-372-96
19	89.5	7.1	249	6	US-10-515-868-1
20	86.5	6.9	176	7	US-11-031-206-86
21	86	6.9	498	6	US-10-453-372-94
22	84.5	6.7	453	6	US-10-453-372-82
23	84.5	6.7	453	6	US-10-453-372-108
24	84.5	6.7	453	6	US-10-453-372-110
25	84.5	6.7	453	6	US-10-453-372-112

26	84	6.7	370	6	US-10-641-678-35	Sequence 35, Appl
27	84	6.7	459	6	US-10-641-678-72	Sequence 72, Appl
28	83.5	6.7	453	6	US-10-453-372-104	Sequence 104, App
29	82.5	6.6	236	6	US-10-527-500-65	Sequence 65, Appl
30	81.5	6.5	4655	6	US-10-995-561-556	Sequence 556, App
31	81	6.5	1332	7	US-11-242-459-9	Sequence 9, Appl
32	80.5	6.4	463	6	US-10-453-372-106	Sequence 106, App
33	80.5	6.4	463	7	US-11-154-673-9	Sequence 9, Appl
34	80.5	6.4	550	7	US-11-154-673-8	Sequence 8, Appl
35	80.5	6.4	625	7	US-11-154-673-7	Sequence 7, Appl
36	79.5	6.3	343	7	US-11-087-099-7967	Sequence 7967, Ap
37	79	6.3	164	7	US-11-031-206-84	Sequence 84, Appl
38	78.5	6.3	161	7	US-11-031-206-96	Sequence 96, Appl
39	77	6.1	206	6	US-10-498-026-80	Sequence 80, Appl
40	76.5	6.1	164	7	US-11-031-206-82	Sequence 82, Appl
41	76.5	6.1	253	6	US-10-467-657-5684	Sequence 5684, Ap
42	76	6.1	453	6	US-10-878-556A-141	Sequence 141, App
43	75.5	6.0	373	7	US-11-087-099-1660	Sequence 1660, Ap
44	75.5	6.0	373	7	US-11-087-099-5200	Sequence 5200, Ap
45	75.5	6.0	486	6	US-10-467-962B-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1  
US-10-527-771-10  
; Sequence 10, Application US/10527771  
; Publication No. US20050271683A1  
; GENERAL INFORMATION:  
; APPLICANT: University Gent  
; TITLE OF INVENTION: Ostertagia vaccine  
; FILE REFERENCE: 2002-015  
; CURRENT APPLICATION NUMBER: US/10/527,771  
; PRIOR FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: US 10/243,319  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 10  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Ostertagia ostertagi  
; US-10-527-771-10

Query Match 100.0%; Score 1253; DB 6; Length 231;  
Best Local Similarity 100.0%; Pred. No. 3.8e-123;  
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSAAVVAVVAVLALFSYABAGFCPCNSLSQSDSARQIFLDFHNDVRNRNALGNGLINWTVN	60
Db	1	MSAAVVAVVAVLALFSYABAGFCPCNSLSQSDSARQIFLDFHNDVRNRNALGNGLINWTVN	60
Qy	61	ADAVIILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAONIARWLYFKDSEETV	120
Db	61	ADAVIILGPAQNMKYKVDWDCNLEEVAAQOIAPCNDPLPINTSLAONIARWLYFKDSEETV	120
Qy	121	LQOVSWYVWSASLGFPMKGTLDQFANQWABPLANIANYNRNVGCAHKICPAQONMNVSC	180
Db	121	LQOVSWYVWSASLGFPMKGTLDQFANQWABPLANIANYNRNVGCAHKICPAQONMNVSC	180
Qy	181	VYGSFKLAPNEVIWQEGKACVCDARPDFCCDNLCTFRDASVRHQCASP	231
Db	181	VYGSFKLAPNEVIWQEGKACVCDARPDFCCDNLCTFRDASVRHQCASP	231

RESULT 2  
US-10-527-771-2  
; Sequence 2, Application US/10527771  
; Publication No. US20050271683A1  
; GENERAL INFORMATION:  
; APPLICANT: University Gent  
; TITLE OF INVENTION: Ostertagia vaccine

```
FILE REFERENCE: 2002-015
CURRENT APPLICATION NUMBER: US/10/527,771
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: US 10/243,319
PRIOR FILING DATE: 2002-09-13
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 236
TYPE: PRT
ORGANISM: Osteragia osteragi
FEATURE:
NAME/KEY: misc_feature
LOCATION: (37)..(37)
OTHER INFORMATION: The xaa at location 37 stands for Lys, or Gln.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121)..(121)
OTHER INFORMATION: The xaa at location 121 stands for Asn, or Thr.
US-10-527-771-2
```

Query Match 50.8%; Score 636.5; DB 6; Length 236;  
Best Local Similarity 52.5%; Pred. No. 6.4e-59;  
Matches 125; Conservative 35; Mismatches 67; Indels 11; Gaps 4;

```
QY 1 MSAVVAV--LALFSYAEAGFCPCNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMT 58
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MQALIGIALYLVLTSTNEAGFCPCADINQTDARKIFLDFHNOVRDIAGASPLNLT 60

QY 59 VVADAV---ILGPAQNMVKVWDNCNLEEVAAQOIAPCNDPLPINTSLAQNIARWLYEKD 114
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 ---GAVQMRNVLGPAKMVRMDNCNLEAKAKAMITPCTPLPIDTSPQNLAQWLFQN 117

QY 115 SEETVLQOVSWYVWSASLGFMKGTKLDFANQMAEPLANIANYRNRKVGCAHKIC--PA 172
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 SQEXEVLQTTPMSWVTASLRNLQDTEANITNWQIRPLSNIAWQNLKVGCAHKVCKEPT 177

QY 173 QQNMVSCVYGSPLAPNEVIWQEGKACVCDARPDSPCCDNLCDTRDAASVRHQQCAS 230
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GTNVVSCAYGGEVLQDNEVWVDKGPCTCMNAYPNSFCNNLCDTIAATLRKQPKCS 235
```

RESULT 3

```
US-11-153-222A-1
Sequence 1, Application US/11153222A
Publication No. US20060039921A1
GENERAL INFORMATION:
APPLICANT: Lustigman, Sara
APPLICANT: MacDonald, Angus J.
TITLE OF INVENTION: Adjuvancy and Immune Potentiating Properties of Natural Products
TITLE OF INVENTION: of Onchocerca Volvulus
FILE REFERENCE: 454-37
CURRENT APPLICATION NUMBER: US/11/153,222A
CURRENT FILING DATE: 2005-06-15
PRIOR APPLICATION NUMBER: 60/580,254
PRIOR FILING DATE: 2004-06-15
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1
LENGTH: 223
TYPE: PRT
ORGANISM: Onchocerca volvulus
US-11-153-222A-1
```

Query Match 10.4%; Score 130.5; DB 7; Length 223;  
Best Local Similarity 23.4%; Pred. No. 3.1e-06;  
Matches 60; Conservative 24; Mismatches 97; Indels 75; Gaps 10;

```
QY 4 AVVAVVLLALFSYAEAGFCPCNSLSQSDSARQIFLDFHNDVRRNIALGNGLINMTV--NAD 62
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 AIVVAV-----TGYNCPGKGLTALERKI-----VGQNNKRRSDLINGKLKRN 52

QY 63 AVILGPAQNMVKVWDNCNLEEVAAQOIAPCNDPLPINTSLAQNIARWLYFKDS---EET 119
```

```
Db 53 GTYMPRGKNMLELRWDCKLE-----SSAQRWANOCIFGHSPROOREG 94

QY 120 VLQOVSWYVWSASLGFMKGTKLDFANQMAEPLANIANYRNR----- 161
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 95 VGENVAYWVSVEGLKKTAGTDAGKSWSELPKL--YENNPNNMTWKVAGQGLHFT 152

QY 162 -----KVGCAHKI--CPAQNMVSCVYGSPLAPNEVIWQEGKAC--VCDARPDSPCC 211
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 QMAGKTYKIGCGVATQCDDGGRTLIVICHYSPGNNMVGCVIYGRGNPCNPCKVDKD----- 208

QY 212 DNLCDTRDAASVRHQC 227
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 ---CYTKKCLSKSLC 221
```

RESULT 4

```
US-11-153-222A-2
Sequence 2, Application US/11153222A
Publication No. US20060039921A1
GENERAL INFORMATION:
APPLICANT: Lustigman, Sara
APPLICANT: MacDonald, Angus J.
TITLE OF INVENTION: Adjuvancy and Immune Potentiating Properties of Natural Products
TITLE OF INVENTION: of Onchocerca Volvulus
FILE REFERENCE: 454-37
CURRENT APPLICATION NUMBER: US/11/153,222A
CURRENT FILING DATE: 2005-06-15
PRIOR APPLICATION NUMBER: 60/580,254
PRIOR FILING DATE: 2004-06-15
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.3
SEQ ID NO 2
LENGTH: 253
TYPE: PRT
ORGANISM: Onchocerca volvulus
US-11-153-222A-2
```

Query Match 10.3%; Score 128.5; DB 7; Length 253;  
Best Local Similarity 22.1%; Pred. No. 5.8e-06;  
Matches 55; Conservative 37; Mismatches 88; Indels 69; Gaps 14;

```
QY 4 AVVAVVLLALFSYAEAGFCPCNSLSQSDSARQIFLDFHNDVRRNIALG--NGLINMTVN 60
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 AIVVAV-----TGYDCYRG-KLTPQYREKIVREHRLRSKLAKGYKNSAGKW--- 55

QY 61 ADAVILGPAQNMVKVWDNCNLEVA---AQOIAPCNDPLPINTSLAQNIARWLYPKDSEE 117
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 -----MPKGNMMEKMDCELEMAQRWADQCVSGNSPKDRGRIGENV---YTORSPT 106

QY 118 ETVLQOVSW-----YVWSASLGFMKGTKLDFANQMAEPLAN--IANY-----RNRK 162
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 SVAVYGTSGIMIALESWVELTRSYKNP-----SNKTSIVANRGVSNFTQLAWGKTYK 161

QY 163 VGC---AHKICPAQNMVSCVYGSPLAPNEVIWQEGKACVCDARPDSPCCDNLCDTRD 219
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 VGCGIATH--CDGKAFVAVCOYNPGNTMGESITYEKGPRCKTD-----RD 205

QY 220 AASVRHQQC 228
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 206 CSS--RKCC 212
```

RESULT 5

```
US-10-515-868-5
Sequence 5, Application US/10515868
Publication No. US20050282729A1
GENERAL INFORMATION:
APPLICANT: Hamilton, David W
APPLICANT: Roberts, Kenneth P
APPLICANT: Ensrud, Kathy M
TITLE OF INVENTION: CRISP POLYPEPTIDES AS CONTRACEPTIVES AND INHIBITORS OF SPERM
CAPACITATION
```



```

; FILE REFERENCE: 110.01860101
; CURRENT APPLICATION NUMBER: US/10/515,868
; CURRENT FILING DATE: 2004-11-24
; PRIOR APPLICATION NUMBER: 60/383,628
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: PCT/US03/16669
; PRIOR FILING DATE: 2003-05-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 5
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-515-868-5

```

```

Query Match          9.5%; Score 119; DB 6; Length 243;
Best Local Similarity 21.6%; Pred. No. 5.4e-05;
Matches 42; Conservative 29; Mismatches 85; Indels 38; Gaps 8;

```

```

QY 41 HNDVRNIALGNLINTVNADAVILGPAQNMKYKWDGNCLEEVAAQIAPC-----ND 94
   ||:||||:|
Db 45 HNELRRQVS-----PRGSNLIKMEWNVQAANAQKMANNCILHSSTE 87
   ||:||||:|
QY 95 PLPINTSLAQNIAKRLYFKDSEETVLQOVSVYWSASLGFMTKGLDQFANQWAEPLAN 154
   ||:||||:|
Db 88 DRKINIKGGENL--YMSTDPSTWRTVIQ--SWYEEENENFYFVGAKPNSAVGHYTO---- 139
   ||:||||:|
QY 155 IANYRNKVCAGAHKICPAQNMVY-----SCVYGSFKLAPNEVIWQEGKACV-CDARPD 208
   ||:||||:|
Db 140 LVWYSSFKVGCYVAYCPNQDTLKYFYVCHYCPMGN-NVMKSTPYHQGTCPASCPCNNCDN 198
   ||:||||:|
QY 209 FCCDNLCDTRDAS 222
   ||:||||:|
Db 199 GLCTNSCDFEDLLS 212
   ||:||||:|

```

```

RESULT 6
US-11-153-222A-3
; Sequence 3, Application US/11153222A
; Publication No. US20060039921A1
; GENERAL INFORMATION:
; APPLICANT: Lustigman, Sara
; APPLICANT: Macdonald, Angus J.
; TITLE OF INVENTION: Adjuvancy and Immune Potentiating Properties of Natural Products
; TITLE OF INVENTION: of Onchocerca Volvulus
; FILE REFERENCE: 454-37
; CURRENT APPLICATION NUMBER: US/11/153,222A
; CURRENT FILING DATE: 2005-06-15
; PRIOR APPLICATION NUMBER: 60/580,254
; PRIOR FILING DATE: 2004-06-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 3
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Onchocerca volvulus
US-11-153-222A-3

```

```

Query Match          9.4%; Score 117.5; DB 7; Length 220;
Best Local Similarity 23.1%; Pred. No. 6.8e-05;
Matches 52; Conservative 27; Mismatches 117; Indels 29; Gaps 9;

```

```

QY 4 AVVAVVLALFSYAEAGFCCNSLSQSDSARQIFLDFHNDVRNIALGNLINTVNADA 63
   ||:||||:|
Db 9 AIIYAV-----TGHDCRG-KLTSLQDITIDEHNKRRSLVKN-----FANKDG 53
   ||:||||:|
QY 64 VILGPAQNMKYKWDGNCLEEVAAQIAPC--NDPLPINTSLAQNIAKRLYFKDSE---E 117
   ||:||||:|
Db 54 NSMPKGNMEMEWDCELEISAQNMADQCFGYSPEQREGVGENITYALGLPKDVEVENT 113
   ||:||||:|
QY 118 ETVLQOVSVYWSASLGFMTKGLDQFANQWAEPLANIANRYRNKVCAGAHKI-CPAQ 173
   ||:||||:|
Db 114 SAALPAIESWTELIRSYRNPNPNKLTSSVASQDVLHFTQMAWGKTHKVGCGIANHCDDG 173
   ||:||||:|

```

```

QY 174 QNMVSVYGSFKLAPNEVIWQEGKACVCDARPDSCDNLCDTR 218
   ||:||||:|
Db 174 EAFIVCHYAPRGNTIGELIYEGSPC-----KVNKHCRTRKCSRK 214
   ||:||||:|

```

```

RESULT 7
US-11-090-439-38
; Sequence 38, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-38

```

```

Query Match          8.9%; Score 111.5; DB 7; Length 258;
Best Local Similarity 22.5%; Pred. No. 0.00035;
Matches 53; Conservative 22; Mismatches 62; Indels 99; Gaps 12;

```

```

QY 27 LSQSDSARQIFLDFHNDVRNIALGNLINTVNADAVILGPAQNMKYKWDGNCLEEVAA 86
   ||:||||:|
Db 62 ISQNDMI--ALDYHNOYR-----GKVFPPANMEYMWDENLAKSAE 102
   ||:||||:|
QY 87 QOIAPC--NDPLPINTSLAQNIAKRLYFKDSEETVLQOVSVYWSASLGFMTKGLD 142
   ||:||||:|
Db 103 AWATCTIWDHGPSYLRFLGQNL-----VTRGRYSILQLVKPWY-----D 144
   ||:||||:|
QY 143 QFANQWAEPLANIANRY-----NRKVCAGAHKICPAQNMVSVY 182
   ||:||||:|
Db 145 E-VKDYAFPPQDCNPRCNRGFCPMCTHYTQMWATSNRIGCAIHTC---QNM---NVM 197
   ||:||||:|
QY 183 GSPKLPAPNEVIWQEGKACVCDARPD-----DSFCCDNL 215
   ||:||||:|
Db 198 GS-----VWRRAVYLVCNYPKGNWIGAPYKVGVCSSCPPSYGSCCTDNL 245
   ||:||||:|

```

```

RESULT 8
US-10-515-868-6
; Sequence 6, Application US/10515868
; Publication No. US20050282729A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, David W
; APPLICANT: Roberts, Kenneth P
; APPLICANT: Ensrud, Kathy M
; TITLE OF INVENTION: CRISP POLYPEPTIDES AS CONTRACEPTIVES AND INHIBITORS OF SPERM
; TITLE OF INVENTION: CAPACITATION
; FILE REFERENCE: 110.01860101
; CURRENT APPLICATION NUMBER: US/10/515,868
; CURRENT FILING DATE: 2004-11-24
; PRIOR APPLICATION NUMBER: 60/383,628
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: PCT/US03/16669
; PRIOR FILING DATE: 2003-05-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 6
; LENGTH: 245
; TYPE: PRT
; ORGANISM: HOMOSAPIEN
US-10-515-868-6

```

Query Match 8.5%; Score 106.5; DB 6; Length 245;  
Best Local Similarity 23.6%; Pred. No. 0.0011;  
Matches 54; Conservative 25; Mismatches 99; Indels 51; Gaps 10;

OY 28 SQSDSARQIFLDFHNDVRNRNALGNGLINWTVNADAVILGPAQNMKYKVDWCNLEEVAAQ 87  
Db 34 TQTQVQREI-VNKNELRAVS-----PPARNMLKMEWKEAANAQK 75  
OY 88 QIAPCN---DPLPINTSL--AQNIARWLKFKDSEETVLQOVSWYVWSASIGFMKGTKL 141  
Db 76 WANQCNRYHSNPKDRTSLKCGEN----LYMSASSWSQALIQSWFDEYNDPFGVGPXT 131  
OY 142 DQFANQWAEPLANIANRYNRKVGCAHKICPAQOQ--NMVWSCVYGSFKLAPNE--VIMQEG 197  
Db 132 ---PNAVGHYTTQVWYSSYLVCGNAYCPNQKVLKYYVCQYCPAGNMANRLYVPYEG 188  
OY 198 KACV-CDARPDSFCCDNLCDTRDAAS-----VRHQCAS 230  
Db 189 APCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLCGHQLVRDSCKAS 237

RESULT 9  
US-11-067-573-2

; Sequence 2, Application US/11067573  
; Publication No. US20050260622A1  
; GENERAL INFORMATION:  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Guiney, Austin  
; APPLICANT: Godowski, Paul  
; APPLICANT: Wood, William  
; APPLICANT: Smith, Victoria  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES  
; FILE REFERENCE: 39780-3430R1C349C1  
; CURRENT APPLICATION NUMBER: US/11/067,573  
; PRIOR FILING DATE: 2005-02-25  
; PRIOR APPLICATION NUMBER: US 10/198,766  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: US 10/052,586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: PCT/US01/06520  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 09/380,138  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: PCT/US99/05028  
; PRIOR FILING DATE: 1999-03-08  
; PRIOR APPLICATION NUMBER: US 60/087,098  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-067-573-2

Query Match 8.3%; Score 103.5; DB 7; Length 500;  
Best Local Similarity 22.2%; Pred. No. 0.0053;  
Matches 50; Conservative 26; Mismatches 68; Indels 81; Gaps 13;

OY 30 SDSARQIFLDFHNDVRNRNALGNGLINWTVNADAVILGPAQNMKYKVDWCNLEEVAAQOI 89  
Db 58 TDNDMQSILDLHNKLR-----SQVYPTASNMEYMTDVELERSAESWA 100  
OY 90 APC--NDPLPINTSLAQNI-ARWLKFKDSEETVLQOVSWYVWSASIGFMKGTKLQDFA 145  
Db 101 ESCLWEHGPAISLPSIGQNLGAHWGRYR---PPTFHVQSWY-----DEVKDFS 145  
OY 146 NQWAEPLANIANR-----NRKVGCAHKICPAQONM-----VV 178  
Db 146 XPYEHECNPYCPFRCSGPVCTHYTQVWVWATSNRIGCAINLC--HNMNIWGQIWPKAVYL 202

OY 179 SCVYGSPE-----LAPNEVIQEGKACVCDARPDSF---CCDNLIC 215  
Db 203 VCNY-SPKGNWVGHP-----YKHGRP--CSACPPSFGGCGRENLC 240

RESULT 10  
US-10-515-868-4

; Sequence 4, Application US/10515868  
; Publication No. US20050282729A1  
; GENERAL INFORMATION:  
; APPLICANT: Hamilton, David W  
; APPLICANT: Roberts, Kenneth P  
; APPLICANT: Ensrud, Kathy M  
; TITLE OF INVENTION: CRISP POLYPEPTIDES AS CONTRACEPTIVES AND INHIBITORS OF SPERM  
; FILE REFERENCE: 110.01860101  
; CURRENT APPLICATION NUMBER: US/10/515,868  
; PRIOR FILING DATE: 2004-11-24  
; PRIOR APPLICATION NUMBER: 60/383,628  
; PRIOR FILING DATE: 2002-05-28  
; PRIOR APPLICATION NUMBER: PCT/US03/16669  
; PRIOR FILING DATE: 2003-05-28  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: HOMOSAPIEN  
US-10-515-868-4

Query Match 7.9%; Score 99; DB 6; Length 243;  
Best Local Similarity 22.8%; Pred. No. 0.0065;  
Matches 47; Conservative 26; Mismatches 97; Indels 36; Gaps 10;

OY 28 SQSDSARQIFLDFHNDVRNRNALGNGLINWTVNADAVILGPAQNMKYKVDWCNLEEVAAQ 87  
Db 32 TQTQVQREI-VNKNELRAVS-----PPASNMLKMEWSREYTTNAQR 73  
OY 88 QIAPC---NDPLPINTS--LAQNIARWLKFKDSEETVLQOVSWYVWSASIGFMKGTKL 141  
Db 74 WANKCTLQHSDEPDRKSTRCGEN---LYM--SSDPTSWSAISQSWYDEILDFTYGVG- 126  
OY 142 DQFANQWAEPLANIANRYNRKVGCAHKICPAQONM--VWSCVY--GSPKLPNEVIWQEG 197  
Db 127 PKSPNAVGHYTTQVWYSTYQVCGGIAYCPNQDSLKYVVCQYCPAGNMNRKNTPTPYQOG 186  
OY 198 KACV-CDARPDSFCCDNLCDTRDAAS 222  
Db 187 TPCAGCPDDCDKGLCTNSCQYQDLLS 212

RESULT 11

US-10-506-443A-41  
; Sequence 41, Application US/10506443A  
; Publication No. US20060013817A1  
; GENERAL INFORMATION:  
; APPLICANT: Sahin Dr., Ugur  
; APPLICANT: Tureci Dr., Ozlem  
; APPLICANT: Koslowski Dr., Michael  
; TITLE OF INVENTION: Genetic Products Differentially Expressed In Tumors and Use There  
; FILE REFERENCE: 342-3PCT  
; CURRENT APPLICATION NUMBER: US/10/506,443A  
; PRIOR FILING DATE: 2004-09-01  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-506-443A-41

Query Match 7.9%; Score 99; DB 6; Length 243;  
Best Local Similarity 22.8%; Pred. No. 0.0065;

Matches 47; Conservative 26; Mismatches 97; Indels 36; Gaps 10;

QY 28 S QSDSARQIFLDFHNDVRNIALGNGLINVTNADAVILGPAQNMKYVDWCNLEEVAQAQ 87  
 Db 32 TOLQVQREI-VNKNELRKAVS-----PPASNMLKNEWSREVTNAQR 73

QY 88 QIAPC---NDPLPINTS--LAQNIARWLYFKDSEETVLQOVSWYWSASLGFMKGTKL 141  
 Db 74 WANKTQLQHSDEPDRKTRCGEN---LYM--SSDPTWSAIGSWYDEILDVYGVG- 126

QY 142 DQFANQWAEPLANIANIYRNKRKVGCAHKICPAQNM--VVSQVY--GSPKLAPEVIMQEG 197  
 Db 127 PKSPNAVGHYTLQVWYSTYQVCGIAYCPNQDSLKYVYVQYCPAGNNMNRKNTPTQCG 186

QY 198 KACV--CDARPDSCCDNLCDTRDAAS 222  
 Db 187 TPCAGCPDDCDKGLCTNSCQYQDLIS 212

RESULT 12

US-10-527-500-23  
 ; Sequence 23, Application US/10527500  
 ; Publication No. US20060004186A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
 ; REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
 ; HUMAN SERVICES  
 ; APPLICANT: Valenzuela, Jesus G.  
 ; APPLICANT: Ribero, Jose M.C.  
 ; APPLICANT: Kamhawi, Shaden  
 ; APPLICANT: Belkaid, Yasmine  
 ; APPLICANT: Fischer, Laurent Bernard  
 ; APPLICANT: Audonnet, Jean-Cristophe  
 ; APPLICANT: Mliward, Francis William  
 ; TITLE OF INVENTION: P. ARIASI POLYPEPTIDES AND P. PERNICIOSUS POLYPEPTIDES AND  
 ; TITLE OF INVENTION: METHODS OF USE  
 ; FILE REFERENCE: 4239-66903-02  
 ; CURRENT APPLICATION NUMBER: US/10/527,500  
 ; PRIOR FILING DATE: 2005-03-11  
 ; PRIOR APPLICATION NUMBER: PCT/US2003/029833  
 ; PRIOR FILING DATE: 2003-09-18  
 ; PRIOR APPLICATION NUMBER: US 60/425,852  
 ; PRIOR FILING DATE: 2002-11-12  
 ; PRIOR APPLICATION NUMBER: US 60/412,327  
 ; PRIOR FILING DATE: 2002-09-19  
 ; NUMBER OF SEQ ID NOS: 87  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 23  
 ; LENGTH: 277  
 ; TYPE: PRT  
 ; ORGANISM: Phlebotomus ariasi  
 ; US-10-527-500-23

Query Match 7.8%; Score 97.5; DB 6; Length 277;

Best Local Similarity 21.1%; Pred. No. 0.011;

Matches 56; Conservative 41; Mismatches 101; Indels 67; Gaps 15;

QY 5 VVAVALLAFSYABAGFC-----CP-----NSLSQSDSARQIF 37  
 Db 11 VVLFVAHSNDYCEPKLCKFNNOVTHIGCKNDGKFVSTCPKPNDAQIMDTQQRKMLF 70

QY 38 LDFHNDVRNIALGNGLINVTNADAVILGPAQNMKYVDWCNLEEVAQAQIAPC--ND 94  
 Db 71 LKIHNRRLRLARGS-VSNFK-----SAAKMPMLKWDNELARLAAYNRVTRCKPAHD 120

QY 95 PLPINTS---LAQNIARWLYPKD-SSEETVLQOVSWYWS---ASLG---FMKGTCLD 142  
 Db 121 QCRSTYACPYAGQNLGQMLSSPFLDPNVYIKITREWFLEVKWANGHTDKYMTG---- 176

QY 143 QFANQWAEPLANIANIYRNKRKVGCA-HKICPAQNM--VVSQVYGSPKLAPEVIMQ 195  
 Db 177 --SGKNGKAIGHFTAFIHEKSDKVGCAVAKLTNOQYNNMKQYLVACNYCYTMTL-KEGIYT 233

QY 196 BGKAC-VCDARPDSCCDNLCDTRD 219  
 Db 234 TGKPCSQCGKKCDSDVYKMLCDASE 258

RESULT 13

US-10-453-372-100  
 ; Sequence 100, Application US/10453372  
 ; Publication No. US20060003323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alsbrook, et al.  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-589 A  
 ; CURRENT APPLICATION NUMBER: US/10/453,372  
 ; PRIOR FILING DATE: 2003-06-03  
 ; PRIOR APPLICATION NUMBER: 09/789390  
 ; PRIOR FILING DATE: 2001-02-23  
 ; PRIOR APPLICATION NUMBER: 60/185967  
 ; PRIOR FILING DATE: 2000-03-01  
 ; PRIOR APPLICATION NUMBER: 09/823187  
 ; PRIOR FILING DATE: 2001-03-29  
 ; PRIOR APPLICATION NUMBER: 60/195792  
 ; PRIOR FILING DATE: 2000-03-10  
 ; PRIOR APPLICATION NUMBER: 09/839446  
 ; PRIOR FILING DATE: 2001-03-19  
 ; PRIOR APPLICATION NUMBER: 60/199476  
 ; PRIOR FILING DATE: 2000-03-25  
 ; PRIOR APPLICATION NUMBER: 09/863776  
 ; PRIOR FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: 60/208263  
 ; PRIOR FILING DATE: 2000-05-31  
 ; PRIOR APPLICATION NUMBER: 09/939398  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: 60/227800  
 ; PRIOR FILING DATE: 2000-08-25  
 ; Remaining Prior Application data removed - See file wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1609  
 ; SOFTWARE: CuraSeqList version 0.1  
 ; SEQ ID NO 100  
 ; LENGTH: 497  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-453-372-100

Query Match 7.4%; Score 92.5; DB 6; Length 497;

Best Local Similarity 21.0%; Pred. No. 0.074;

Matches 50; Conservative 28; Mismatches 91; Indels 69; Gaps 12;

QY 10 LIALFSYABAGFCPSNLSQSDSARQIFLDFHNDVRNIALGNGLINVTNADAVILGPA 69  
 Db 34 LLSKYQHNEHSRVRRAIPRED--KEEILMLHNKLRGVQ-----PQA 74

QY 70 QNMKYVDWCNLEEVAQAQIAPC---NDPLPINTSLAONI-ARWLYFKDSEETVLQOVS 125  
 Db 75 SNMEYMTWDELKSAAMAWASQCIWEHGPTSLVLSIGQNLGAHWGRYRSPG----FHVQS 130

QY 126 WY-----WVSASLGFMKGTCLDQFANQWAEPLANIANIYRNKRKVGCAHKI 169  
 Db 131 WYDEVKDYTPYPSECNPCBERCSGPMCTHYTQIV--WA-----TTNKIGCAVNT 179

QY 170 CPAQNMV-----SCVYGSPKLAPE-----NEVIMQEGKAC-VCDARPDSCCDNL 215  
 Db 180 C---RKMTVGEVWENAVYFVCNYSFGKNGWIGEAIFYKNGRCPSCPPSYGSGCRNNLC 234

RESULT 14

US-10-453-372-102  
 ; Sequence 102, Application US/10453372  
 ; Publication No. US20060003323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alsbrook, et al.  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-589 A



```

; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Cursesqlst version 0.1
; SEQ ID NO 102
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-102

```

Query Match	7.4%	Score 92.5;	DB 6;	Length 497;
Best Local Similarity	21.0%;	Pred. No. 0.074;		
Matches 50;	Conservative 28;	Mismatches 91;	Indels 69;	Gaps 12;

```

QY      10 LIALFSYAEGFCCPNLSQSDSARQIFLDFHNDVRNIALGNGLINWTVANDAVILGPA 69
      ||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      34 LISKQHNESHSHRVRRRAIPRED--KEEILMLHNKLRGQVQ-----PQA 74

QY      70 QNMVKVDWCNLEEVAAQOIAPC--NDPLPINTSLAQNI-ARWLTFKDSSEETVLQQVS 125
      ||: : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db      75 SNMEYMTWDELEKSAAMASQCIWEHGPTSLVLSIGQLGAHWGRYRSPG----FHVQS 130

QY      126 WY-----WVSASLGFMTGKTDQFANQWAEPLANIANRYNRKVGCAHKI 169
      |||||
Db      131 WYDEVKDYTYYPYSECNPWCPCERCSGPMCTHYTQIV--WA-----TTNKIGCAVNT 179

QY      170 CPAQQNMVV-----SCVYGSPTLAP-----NEVIWQEGKAC-VCDARPDSPCCDNLIC 215
      ||: : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db      180 C---RKMTVWGEVWENAVTFVVCNYSPKGNWIGELAPYKNGRPCSECPSPSYGSCANNLIC 234

```

RESULT 15  
US-10-453

US-10-453-372-98

; Sequence 98, Application US/10453372  
; Publication No. US2006000323A1

GENERAL INFORMATION:

APPLICANT: Albobrook, et al.

1. TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

FILE REFERENCE: 21402-589 A

CURRENT APPLICATION NUMBER: US/10/453,372 ;

**CURRENT FILING DATE:** 2003-06-03

PRIOR APPLICATION NUMBER: 09/7899

PRIOR FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: 60/185967

PRIOR FILING DATE: 2000-03-01  
PRIORITY ADDITION NUMBER: 00/69331987

PRIOR APPLICATION NUMBER: 09/823187  
 PRIOR FILING DATE: 2001-03-29

PRIOR FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 60/198579

PRIOR APPLICATION NUMBER: 60/195192

PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 09/839446

PRIOR FILING DATE: 2001-03-19

PRIOR APPLICATION NUMBER: 60/199476

PRIOR APPLICATION NUMBER: 80/153418

```

: PRIOR FILING DATE: 2000-03-25
: PRIOR APPLICATION NUMBER: 09/863776
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: 60/208263
: PRIOR FILING DATE: 2000-05-31
: PRIOR APPLICATION NUMBER: 09/939398
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: 60/227800
: PRIOR FILING DATE: 2000-08-25
: Remaining Prior Application data removed - See File Wrapper or PALM
: . NUMBER OF SEQ ID NOS: 1609
: SOFTWARE: Curaseq1st version 0.1
: SEQ ID NO 98
: LENGTH: 475
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-453-372-98

```

Query Match	7.3%;	Score 91.5;	DB 6;	Length 475;
Best Local Similarity	21.0%;	Pred. No. 0.089;		
Matches 50; Conservative	28;	Mismatches 91;	Indels 69;	Gaps 12;

Qy		10	L L A F S Y A E A G F C C P N S L S O S D S A R Q I F L D H N D V R R N I A L G N G L I N T W A N A D A V I L G P A	69
			: : : :   : : :     : :	
Db		12	L L S K Q J H N E S H S R V R R A I P R E D - K E E I I M L H N K L R G Q V O - - - - -	P Q A 52
Qy		70	Q N N Y K V D W D C N L E E V A A Q Q I A P C - - N D P L P I N T S L A O N I - A R M L Y F K D S E E T V L Q O V S	125
			: :     :   :   : :     :   :   : :	
Db		53	S N M E Y W T W D D E L E K S A A A W A S Q C I W E H G P T G L L V S I G O N L G A H W G R Y R S P G - - - F H V Q S	108
Qy		126	W Y - - - - - - - - - - W S A S L G F M K G T K L D Q F A N Q W A B P L A N I A N Y R N R K V G C A H K I	169
			: :     :   :   :   :   :   :	
Db		109	W Y D E V X D Y T Y P Y P S E C N P M C P E R C S G P M C H T Y T Q I V - W A - - - - - T T N K I G C A V N T	157
Qy		170	C P A Q Q N N V - - - - - S C V Y G S P K L A P - - - - - N E V I W Q E G K A C - V C D A R P D S F C C D N L C	215
			: : : : : : : : : : : : : : : : : :	
Db		158	C - - R O M T W G E V M E N A V I V F V C N Y S P K G N I T G E A P Y K N G R P C S E C P S Y G S C R N N L C	212

Search completed: March 3, 2006, 19:14:24  
Job time : 22 secs

Job time : 22 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:05:52 ; Search time 40 Seconds  
(without alignments)  
555.652 Million cell updates/sec

Title: US-10-527-771-10  
Perfect score: 1253  
Sequence: 1 MSAVVAVLALPSYAEAG.....DNLCDTRDAASVRHCCASP 231

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*  
1: pirl:\*  
2: pirl2:\*  
3: pirl3:\*  
4: pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	16.8	425	2 C89753	protein F11C7.3 (l
2	178.5	14.2	221	2 T24494	hypothetical prote
3	167	13.3	196	2 T27833	hypothetical prote
4	160	12.8	241	2 T33397	hypothetical prote
5	146	11.7	212	2 T27834	hypothetical prote
6	142.5	11.4	207	2 T22438	hypothetical prote
7	135.5	10.8	213	2 T22439	hypothetical prote
8	121	9.7	207	2 T22432	hypothetical prote
9	120	9.6	208	2 T19852	hypothetical prote
10	119.5	9.5	207	2 T31959	hypothetical prote
11	119	9.5	211	2 T19859	hypothetical prote
12	119	9.5	243	2 JE0204	testicular protein
13	117	9.3	210	2 T19849	hypothetical prote
14	116.5	9.3	212	2 T22437	hypothetical prote
15	115.5	9.2	246	2 T24493	hypothetical prote
16	113.5	9.1	243	2 A33329	testis-specific pr
17	113	9.0	274	2 A54419	neutrophil inhibit
18	111.5	8.9	209	2 T19847	hypothetical prote
19	109.5	8.7	209	2 T19848	hypothetical prote
20	107.5	8.6	232	2 T33704	hypothetical prote
21	105.5	8.4	245	2 S68691	neutrophil granule
22	103	8.2	385	2 T21763	hypothetical prote
23	99	7.9	243	2 B33329	cysteine-rich secr
24	95.5	7.6	208	2 T20661	hypothetical prote
25	92	7.3	266	2 JC5308	testis-specific, v
26	89.5	7.1	249	2 S68684	cysteine-rich secr
27	89	7.1	219	2 JC4131	glioma pathogene
28	88	7.0	464	2 JC7143	endoglucanase I -
29	87.5	7.0	207	2 T22436	hypothetical prote

30	84.5	6.7	212	2 B37330	venom allergen III
31	84	6.7	459	2 A25928	cellulase (EC 3.2.
32	83.5	6.7	164	2 T04299	pathogenesis-relat
33	83.5	6.7	198	2 T26861	hypothetical prote
34	83	6.6	211	2 B58853	venom allergen Sol
35	82.5	6.6	164	2 S52627	pathogenesis-relat
36	81.5	6.5	164	2 S39474	pathogenesis-relat
37	81	6.5	168	2 JC7330	acidic pathogene
38	81	6.5	453	1 A32629	ubiquinol-cytochro
39	80.5	6.4	178	2 S68681	cysteine-rich secr
40	80.5	6.4	647	2 E82579	acetyl coenzyme A
41	80.5	6.4	914	2 C69284	surface layer prot
42	80	6.4	504	2 A83286	acetyl-CoA C-acety
43	80	6.4	737	2 T16737	hypothetical prote
44	80	6.4	2473	1 S38040	1-phosphatidylinos
45	80	6.4	4543	1 A53102	alpha-2-macroglobu

ALIGNMENTS

RESULT 1

C89753  
protein F11C7.3 [imported] - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C/Accession: C89753  
R/Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A/Reference number: A75000, MUID:99069613, PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elc  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A/Accession: C89753  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-425 <STO>  
A/Cross-references: UNIPROT:Q19348; UNIPARC:UPI0000075DA3; GB:chr\_X; PIND:AA69015.1; PI  
C/Genetics:  
A/Gene: F11C7.3.  
A/Map position: X

Query Match	16.8%;	Score 211;	DB 2;	Length 425;
Best Local Similarity	26.5%;	Pred. No. 6.8e-11;		
Matches	65;	Conservative	36;	Mismatches 76; Indels 68; Gaps 10
QY	3	AAVVAVALLALFSYAEA-GFCCPNSLSQSDSARQIFLDFHNDVRNIALG----	NGLIN	56
DB	2	AVLAVVLLACLERAVALQTFGCSNT-KINDQARKMFDANDARRSMAGLEPNKCGL--		58
QY	57	WTVNADAVILGPAQNMKYVDWDCNLEBVAQOIAPC-----	ND	94
DB	59	-----LSGKKNVYELNWDCEMAKAQEWADGCPSSFTQTFDPTWGQNYATYMGSIAD		109
QY	95	PLPINTSLAQNIARWLVPKDSSEETVLQOVSWYVWSASLGFMKGTKLDQFANQWAEPLAN		154
DB	110	PLPY-ASMAVN-----GWMSEIRTVGLTDPD--NKYTNAMFRFRAN		147
QY	155	IANYRNKVGCAHKICPAQNMVVSQVYSPKLAAPNEVIMQEGKACVCDAR----	PDSFC	210
DB	148	MANGKASAFGCAYALCAGK--LSINCITNKIGYMTAIIYEKGDACTSDAECTYSDSOC		205
QY	211	CDNLC	215	
DB	206	KNGLC	210	

RESULT 2  
T24494  
hypothetical protein T05A10.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T24494

R,Sulston, J.  
submitted to the EMBL Data Library, November 1995  
A,Reference number: Z19898  
A,Accession: T24494  
A,Status: preliminary; translated from GB/EMBL/DBJ  
A,Molecule type: DNA  
A,Residues: 1-221 <WIL>  
A,Cross-references: UNIPROT:P90959; UNIPARC:UPI00008266B; EMBL:Z68108; PIDN:CAA92137.1;  
A,Experimental source: clone T05A10  
C,Genetics:  
A,Gene: CESP:T05A10.4  
A,Map position: X  
A,Introns: 18/3; 54/1; 106/3; 142/2; 190/3

	Query Match	14.2%;	Score 178.5;	DB 2;	Length 221;
	Best Local Similarity	29.3%;	Pred. No. 2.4e-08;		
	Matches	51;	Conservative	35;	Mismatches 59; Indels 29; Gaps 8.
QY	22 CCENSLSQSDSARQIFLDFHNDVRRNIALG-----NGLINWTVNADAVILGPAQNMYKVD	76			
Db	61 CNKSTITQLQ--QEILTTHNELRRLRSIAFGQRNKRGIMN-----GARMMYKLD	107			
QY	77 WDCNLEEVAAGQIAPCND---PLPINTSLAQNIRWLIFYKDSSEETV-LQQVSWTWVSAS	132			
Db	108 WDCELASLANWSTSCPQHFMPOSVLGSNAQLFKRFYFYFDGHDSFTVMRNAMKYWWQQ-	166			
QY	133 LGFHKGTKLDQ----PANQMAEPLANTANYRNKKVGCAHKICPAQQNNVYSCVY	182			
Db	167 -GREEKGNE-DQGRFRFYARRNYFGWMANAKGTYTRVGCYSIMCGDGESALFTCLY	218			

### RESULT 3

12/833  
hypothetical protein ZK384.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T27833  
R:Ainscough, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20427  
A:Accession: T27833  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-196 <WIL>  
A:Cross-references: UNIPROT:O62507; UNIPARC:UPI000007E62C; EMBL:Z82092; PIDN:CAB05010.1;  
A:Experimental source: clone ZK384  
C:Genetics:  
A:Gene: CESP:ZK384.1  
A:Map position: 5  
A:Introns: 58/3; 95/3; 157/2  
C:Superfamily: yellowjacket venom allergen antigen 5

	Query Match	13.3%;	Score 167;	DB 2;	Length 196;	
	Best Local Similarity	26.1%;	Pred. No. 2.le-07;			
	Matches	63;	Conservative	30;	Mismatches	70;
					Indels	78;
					Gaps	12;
Qy	5 VVAVALLLESYAEAGFCCPNSLSQSBSARQIFLDFHNDVRNRNALGNGLINMTVNADAV	:	:	:	:	:
Dd	3 LLLFLLAITS--SSGQLSPN-----GRQQVLDFHNKLRSGVALG-----VFSA NGT	:	:	:	:	:
Qy	65 ILGPAQNMYKVWDMDCNLEEVA AQOIAPCNDPL--PINTSLAONIA RMLYFKDSEETVLQ	:	:	:	:	:
Dd	48 IKPPARNMERLTYYGQGFERLAQDYVADCDEGLEIPIGRNIGMN-----	:	:	:	:	:
Qy	123 QVSWYWVSASILGFMKGTKLDQFANOWAEP-----LANIANYRN-----RKVG	:	:	:	:	:
Dd	91 ----YYT-----TKVIDALNDWAEEFQVNGWLSTIYNNTSISASAQMVWAGTKYVG	:	:	:	:	:
Qy	165 CAHKICPAQNNMNVSVCYGSPKLAPNEVIWGEKAC-----VCDA RPDSFCDD---NL	:	:	:	:	:
Dd	138 CGVKRCD-PINVVVVCMYTYQGGLVGRPIYKEGP PCTACPMRICPGQKE--CCDRVMGL	:	:	:	:	:
Qy	215 C 215					

Db 195 C 195

## RESULT 4

135577  
 hypothetical protein B0545.3 - *Caenorhabditis elegans*  
 C/Species: *Caenorhabditis elegans*  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T33397  
 R/Rohlfing, T.  
 submitted to the EMBL Data Library, July 1998  
 A/Description: The sequence of *C. elegans* cosmid B0545.  
 A/Reference number: Z21337  
 A/Accession: T33397  
 A/Status: preliminary; translated from GB/EMBL/DBDJ  
 A/Status: preliminary; translated from GB/EMBL/DBDJ  
 A/Molecule type: DNA  
 A/Residues: 1-241 <ROH>  
 A/Cross-references: UNIPROT:O76663; UNIPARC:UPI000007723F; EMBL:AF078781; PIDN:AAC26915.  
 A/Experimental source: strain Bristol N2; clone B0545  
 C/Genetics:  
 A/Gene: CESP.B0545.3  
 A/Map position: 4  
 A/Introns: 38/3; 68/3; 123/3; 155/3; 194/2

Query Match	12.8%	Score 160	DB 2	length 241
Best Local Similarity	29.0%	Pred. No. 1.le-06		
Matches	47	Conservative	27	Mismatches 50; Indels 38; Gaps 8;
QY	34	RQIFLDFHNDVRNRNIALGNGLINWTVNADAVILGPAQNNKYVDWDNCNLEBVAQAQIAPCN	93	
		: :         : :   :                 : :		
Db	33	QNLLDKHNEIRSQVALG---QYAVDDD--YLPAPADNMVKLDWDCELELEAQQRAQQCN	86	
QY	94	DELPLNTSLAQNIARW-----LYFKDSE---EETVLQQVSWYVWSASLGFMKGTKL	141	
		: :   : :   : :   : :   : :   : :   : :		
Db	87	--LQKENS--GRQMGWDEVRGENAFYFRTTDGLDVSGAVLKGIQRMGDEIAIAGIKNLKL	143	
QY	142	DQFANQ-----WAEPLANIANRYNRKRVGCAHKICPAQON	175	
		: : : :   : :   : :   : :   : :   : :		
Db	144	SRYDSRIGHATQILWKE-----TRKLGCAVQECPPARQD	176	

## RESULT 5

12/834  
hypothetical protein ZK384.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T27834  
R:Ainscough, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20427  
A:Accession: T27834  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-212 <WIL>  
A:Cross-references: UNIPROT:O62508, UNIPARC:UPI000007591B, EMBL:Z82092; PIDN:CAB05011.1;  
A:Experimental source: clone ZK384  
C:Genetics:  
A:Gene: CESP:ZK384.2  
A:Map position: 5  
A:Introns: 60/3; 106/3; 180/2  
C:Superfamily: yellowjacket venom allergen antigen 5

[illegible]



Db 66 LAGMAKONNAATCPSLFTDSKMLGRNYHRLANTSGSLDKYALPAVKKWERQFOERGW-- 123  
 Qy 129 VSASLGFEMKGTLDQFANQ-----WAEPLANIANYNRKYCAHKICPAQONM-- 176  
 Db 124 -----KNOEFRMFGDHRLLTSATQWVA-----TTRHVGCGVNICDAEKNLFG 166  
 Qy 177 -----VSCVYGSPKLAPENEVIWQEGKACVCDARPDSPCCD---NLG 215  
 Db 167 YRNKVAVICEYQSKGNHGLPIYKSGP--TCSACPASTKCERRSGLC 211

RESULT 6  
 T22438

hypothetical protein F49E11.6 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T22438  
 R/Baynes, C.  
 submitted to the EMBL Data Library, March 1996  
 A/Reference number: Z19564  
 A/Accession: T22438  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-207 <WIL>  
 A/Cross-references: UNIPROT:Q93746; UNIPARC:UPI000007498; EMBL:Z70308; PIDN:CAA94350.1;  
 A/Experimental source: clone F49E11  
 C/Genetics:  
 A/Gene: CESP:F49E11.6  
 A/Map position: 4  
 A/Introns: 58/3; 104/3; 176/2  
 C/Superfamily: yellowjacket venom allergen antigen 5

Query Match 11.4%; Score 142.5; DB 2; Length 207;  
 Best Local Similarity 23.6%; Pred. No. 3.2e-05;  
 Matches 56; Conservative 19; Mismatches 79; Indels 83; Gaps 12;  
 Qy 22 CPEPNSLSQ-SDSARQIFLDFHNDVRNIALGNGLINMTVNADAVILGPAQNMKYKWDON 80  
 Db 10 CIAGVFSQFTSTGQAIYDAHNKLRSSIAKGTVAAGTTQKS-----GSMRKIKWDAT 63  
 Qy 81 LEEVAAQOIAPCNDPLPINTSLAQNIAKWLKFKDSEETVLQOVSWYVVSASLGFEMKGT 140  
 Db 64 VATSAQNYANTC---PTGHSQSGSYGENLY-----WYWTSGTIG-----N 100  
 Qy 141 LDQF-----ANQ--WAEPLANIANYNRKYGC 165  
 Db 101 LDTFGPAASSSWSEFPQGYGWTNTLDMNTFNTGIGHATQMAWANTFA-----IGC 151  
 Qy 166 AHKIC---PAQ-QNMVVSVCYGSPLAPNEVIWQEGKACVCDARPDSPCCDN--LC 215  
 Db 152 GYKNGCKDPSNGYNKVAVVCQYKTPGNYLNPYQGGTTCA--ACPSGTACDSSGLC 206

RESULT 7  
 T22439

hypothetical protein F49E11.4 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T22439  
 R/Baynes, C.  
 submitted to the EMBL Data Library, March 1996  
 A/Reference number: Z19564  
 A/Accession: T22439  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-213 <WIL>  
 A/Cross-references: UNIPROT:Q93747; UNIPARC:UPI000007494FD; EMBL:Z70308; PIDN:CAA94351.1;  
 A/Experimental source: clone F49E11  
 C/Genetics:  
 A/Gene: CESP:F49E11.4  
 A/Map position: 4  
 A/Introns: 14/1; 60/3; 105/3; 175/2  
 C/Superfamily: yellowjacket venom allergen antigen 5

Query Match 10.8%; Score 135.5; DB 2; Length 213;  
 Best Local Similarity 25.2%; Pred. No. 0.00014;  
 Matches 59; Conservative 26; Mismatches 92; Indels 57; Gaps 12;

Qy 9 VLLALFSYABAGFCCPNSLSQSDSARQIFLDFHNDVRNIALGNGLINMTVNADAVILGP 68  
 Db 4 VLLILFASA-IGWSDNF---SKGQNLNLNVHNEFRQALGQ-----LSFRGYKPS 53  
 Qy 69 AQNMKYKWDONLEEVAAQOIAPCNDPLPINTSLAQNIAKWLKFKDS-----EEETVLQ 122  
 Db 54 ASMRKISWSKLTJNATKPAETC---PKHNVVMTGESIFWHFSSSLSTPEQYATLA 109  
 Qy 123 QVSWYVVSASLGFEMKGTLDQF-ANQWAEPLANIANYNR-----NRKYCAHK 168  
 Db 110 PQKWW-----NEEFTNGWDSLTYNHASQREPIGHAQVQMAWHTTSKYGCGYS 155  
 Qy 169 ICPA---QNMVVSVCYGSPLAPNEVIWQEGKACV-CDARPDSP--CCDNLCD 216  
 Db 156 KCAVGTPEQTMVVCVRFQKGNIEGEPITYNEGFTCTKC---PEEYQKCPSGLC 206

RESULT 8  
 T22432

hypothetical protein F49E11.10 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T22432  
 R/Baynes, C.  
 submitted to the EMBL Data Library, March 1996  
 A/Reference number: Z19564  
 A/Accession: T22432  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-207 <WIL>  
 A/Cross-references: UNIPROT:Q20603; UNIPARC:UPI000007D17A; EMBL:Z70308; PIDN:CAA94344.1;  
 A/Experimental source: clone F49E11  
 C/Genetics:  
 A/Gene: CESP:F49E11.10  
 A/Map position: 4  
 A/Introns: 58/3; 104/3; 175/2  
 C/Superfamily: yellowjacket venom allergen antigen 5

Query Match 9.7%; Score 121; DB 2; Length 207;  
 Best Local Similarity 23.0%; Pred. No. 0.0025;  
 Matches 56; Conservative 25; Mismatches 89; Indels 74; Gaps 10;  
 Qy 5 VVAVVLLALFSYABAGFCCPNSLSQSDSARQIFLDFHNDVRNIALGNGLINMTVNADAV 64  
 Db 4 LLLVVALAVGCSADFG-----SSQNGIINAHNTLSKIAKG-----TYVAKGT 47  
 Qy 65 ILGPAQNMKYKWDONLEEVAAQOIAPCNDPLPINTSLAQNIAKWLKFKDSEETVLQOV 124  
 Db 48 QKSPGTNLKMKWD---SAVAAS-----AQNYANGCPTGHSGDAGLGENTL 89  
 Qy 125 SWYVVSASLGFEMKGTLDQFANQ-----WAEPLANI-----ANY 158  
 Db 90 YMYWTSGSLG-----DLNQYGSASASWEKEFDYGWKSNLMTIDLFTGTGIGHATQMAWA 144  
 Qy 159 RNRKYCAHKICPAQON---MVVSCVYGSPKLAPENEVIWQEGKACVCDARPDSPCCD-- 212  
 Db 145 KSNLIGCGVDCGRDSNGLNKVTVCQYKPGQNFINGIYVSG--ATCSGCPSGTSCETS 202  
 Qy 213 -NLG 215  
 Db 203 TGLC 206

RESULT 9  
 T19852

hypothetical protein C39B9.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004



```
C;Accession: T19852  
R;Stms, M.  
submitted to the EMBL Data Library, March 1996  
A;Reference number: Z19187  
A;Accession: T19852  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-208 <WIL>  
A;Cross-references: UNIPROT:O18543; UNIPARC:UPI00000803C8; EMBL:Z70307; PIDN:CAA94335.1;  
A;Experimental source: clone C39E9  
C;Genetics:  
A;Gene: CESP:C39E9.2  
A;Map position: 4  
A;Introns: 14/1; 59/3; 176/2  
C;Superfamily: yellowjacket venom allergen antigen 5
```

Query Match                9.6%; Score 120; DB 2; Length 208;  
Best Local Similarity     24.6%; Pred. No. 0.003;  
Matches     59; Conservative     22; Mismatches     91; Indels     68; Gaps     10;

```
QY      7 VAVLLALFSYAAGFCPCPSLSQSDSARQIFLDHFNDVRRNIALGGLINWTVNADAVIL 66  
        :|::||::||::||::||::||::||::||::||::||::||::||  
DB     5 ILLITLACAGAAYAQF-----SANGQAAILNVHNTLRSLAKG-----TYVAKGTAK 50  
  
QY     67 GPAQNMYKVDWDNCNLLEVAQAQIAPCNDFLPINTSLAONIARWLTFKDSSEETVLQQVSW 126  
        |::||::||::||::||::||::||::||::||::||::||  
DB    51 PAASDMLKKMKWDAT---VAASAQAYAN-KCPTHSGGAAGLGENTLY-----W 92  
  
QY    127 YWVSASLGFMKGTKLDQF-----ANQWAEPPLAN-----IANRYRN 160  
        ||::||::||::||::||::||::||::||::||::||::||  
DB    93 YWTSATY----TNIDQFGATGSAAWEKEFDYDGWSNTLSMSLFNTGIGHATOMAWAKT 147  
  
QY    161 RKVGCAHKICPAQN---MWVCVYGSPKLAPNEVTIQEGKAC-VCDARPDECCDNLC 215  
        ::||::||::||::||::||::||::||::||::||::||  
DB   148 NLIGCGVKNGCKDTNGFNKVTVCOYKPQGNYLNQNIYTSGTTCSKCPSGTSCEAATGLC 207
```

RESULT 10  
T31959  
hypothetical protein F02E11.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T31959  
R;Favella, A.; Scheet, P.  
submitted to the EMBL Data Library, July 1997  
A;Description: The sequence of C. elegans cosmid F02E11.  
A;Reference number: Z21104  
A;Accession: T31959  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-207 <FAV>  
A;Cross-references: UNIPROT:O16575; UNIPARC:UPI00000776B8; EMBL:AF016661; PIDN:AAB66052  
A;Experimental source: strain Bristol N2; clone F02E11  
C;Genetics:  
A;Gene: CESP:F02E11.5  
A;Map position: 2  
A;Introns: 59/3; 175/2  
C;Superfamily: yellowjacket venom allergen antigen 5

Query Match                9.5%; Score 119.5; DB 2; Length 207;  
Best Local Similarity     23.1%; Pred. No. 0.0034;  
Matches     56; Conservative     31; Mismatches     92; Indels     63; Gaps     12;

```
QY      1 MSAVVVAVALFLFSYAAGFCPCPSLSQSDSARQIFLDHFNDVRRNIALGGLINMTVN 60  
        |::||::||::||::||::||::||::||::||::||::||  
DB     1 MRALLTLAVVASVYGGF-----SRAGQKAIVDAHNTLRSSIAKGYVANXTRK 50  
  
QY     61 ADAVILGPAQNMYKVDWDNCNLLEVAQAQIAPCNDFLPINTSLAONTA-RWLYFKDSE 116  
        |::||::||::||::||::||::||::||::||::||::||  
DB    51 E-----PGSNILKKMKWDPTIAKSQAQAYANTC---PTHGSKSYGENLYWRW----- 93  
  
QY   117 EETVLQQVSMYWVSASLGF-----MGKTQLDQ-----FANO--WAEPLANIANY 158  
        |:|::||::||::||::||::||::||::||::||::||::||
```

Db		94	SGAVIKSIDDYGVRRASGAWASEFOKYGWKTNKLDSALFKTGIGHATQMAMAS-----	145
OY		159	RNRKVGCAHKICPAQON----MVVSCVYGSPKLAPNEVIWQEGKAC-VCDARPDSECCDN	213
			:   :: : : : : : : : :	
Db		146	-TGSIGCGVKNCGMIDKNMKYKVAVCQYSARGNMINNITYAGKTCSPAKTKCEKATG	204
OY		214	LC 215	
Db		205	LC 206	

RESULT 11  
T19859  
hypothetical protein F49E11.11 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19859; T22441  
R;Sims, M.  
submitted to the EMBL Data Library, March 1996  
A;Reference number: Z19187  
A;Accession: T19859  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-211 <WIL>  
A;Cross-references: UNIPROT:Q18549; UNIPARC:UPI000007A8CD; EMBL:Z70307; PIDN:CAA94341.1;  
A;Experimental source: clone C39E9  
R;Baynes, C.  
submitted to the EMBL Data Library, March 1996  
A;Reference number: Z19564  
A;Accession: T22441  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-211 <WI2>  
A;Cross-references: UNIPARC:UPI000007A8CD; EMBL:Z70308; PIDN:CAA94354.1; GSPDB:GN00022;  
A;Experimental source: clone F49E11  
C;Genetics:  
A;Gene: CESP:F49E11.11  
A;Map position: 4  
A;Introns: 15/1; 60/3; 179/2  
C;Superfamily: yellowjacket venom allergen antigen 5

Query Match            9.5%; Score 119; DB 2; Length 211;  
Best Local Similarity    25.1%; Pred. No. 0.0038;  
Matches    57; Conservative    25; Mismatches    107; Indels    38; Gaps    10;

OY		5	VVVALLLALFSYAEAGFCPCPSLSQSDSARQIFLDFHNDVRRIALGNGLINVTYNADAV	64
			:::   :: : : : : : : : :	
Db		6	LIVLVTLQIGAYAQP-----RESTQQFIVDLHNKLRISIAKGTYYAKGTTKA---	52
OY		65	ILGPQNMYKVDWDNCLEEVAAQOIAPCNDPPIINTSLAQNTA-----RW--LYEKDSEE	117
			: : : : : : : : : : : : : :	
Db		53	--AGSNLLKKMDTTL-AVTAQTFA---NTCPRGHSNAAGVENLYWRWSLPFGMDI	105
OY		118	ETVLQOVSWMVWSASLGFMGKGTKLDQFANQMAEPLANIANYRNKVGCAHKIC---PAQC	174
			: : : : : : : : : : : : : :	
Db		106	YCGAASVAAWEQEFQYQGWTTNTFTQALANTGIGHATQMAMANTGLIGCVKNCGBPDELN	165
OY		175	N---MVVSCVYGSPKLAPNEVIWQEGKACVCDARPDSECCD---NLG	215
			: : : : : : : : : : : : :	
Db		166	NYNRAVVVCQYKAQGNYLGQDIYKSG--TTCSACPFTGTTCBAATGLC	210

RESULT 12  
JE0204  
testicular protein Tpx-1 - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C;Accession: JE0204  
R;Maeda, T.; Sakashita, M.; Ohba, Y.; Nakanishi, Y.  
Biochem. Biophys. Res. Commun. 248, 140-146, 1998  
A;Title: Molecular cloning of the rat tpx-1 responsible for the interaction between spe  
A;Reference number: JE0204; MUID:98340864; PMID:9675100  
A;Accession: JE0204



OY 80 NLEEVAAQOIAPC-----NDPLP---INTSLAQNIARWLYFKD 114  
DB 97 MLEKFAQNMANNCVFAHSAHYERPNOGONLYMSFSNPDPRLIHTA---VEKW--WQE 150  
OY 115 SEE-----ETVLOQVSWYVWSASLGFMKGTCLDQFANQWAEPLANIANYNRNRKVGAHKI 169  
DB 151 LEEFGTPIDNVLTPELW-----DLKGAIGHY-----TQMAWDRTRYRLGCGIAN 194  
OY 170 CPAQONMVVSCVYGSPLKLPNEVIWQEGKAC-VCDAKPDSECCD---NLC 215  
DB 195 CPKMSYVW--CHYGPAGNRKNKNIYEIGDPCEVDDDCPIGTDCCKTSLC 242

Search completed: March 3, 2006, 19:10:13  
Job time : 41 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2006, 07:04:30 ; Search time 4211 Seconds  
(without alignments)  
10299.585 Million cell updates/sec

Title: US-10-527-771-9

Perfect score: 763  
Sequence: 1 gaggaactgtatgtcgcgcg.....tcattgcaaaaaaaaaa 763

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_ets:\*  
11: gb\_ey:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	763	100.0	763	2 OOS515523	AJ515523 Ostertagi
2	763	100.0	763	6 CQ794479	CQ794479 Sequence
3	273.4	35.8	828	6 CQ794471	CQ794471 Sequence
4	263.8	34.6	712	2 OOS310812	AJ310812 Ostertagi
5	41.4	5.4	2000	6 AX655393	AX655393 Sequence
6	41.4	5.4	110000	15 AP008208_049	Continuation (50 o
7	41.4	5.4	132729	15 AP005849	AP005849 Oryza sat
8	41.2	5.4	187110	9 AC131303	AC131303 Mus muscu
9	41.2	5.4	198900	9 AC123556	AC123556 Mus muscu
10	40.2	5.3	174304	8 AL450327	AL450327 Human DNA
11	39.2	5.1	7218	6 I66494	I66494 Sequence 14
12	39	5.1	95097	8 AC022202	AC022202 Homo sapi
13	38.6	5.1	110000	14 AC121380_2	Continuation (3 of
14	38.6	5.1	225475	14 AC120958	AC120958 Rattus no
15	38.6	5.1	239001	14 AC121042	AC121042 Rattus no
16	38.2	5.0	244038	14 AC115336	AC115336 Rattus no
17	38.2	5.0	254388	14 AC109427	AC109427 Rattus no
18	38.2	5.0	327507	14 AC096293	AC096293 Rattus no

C	19	37.4	4.9	2000	6 AX655393	AX655393 Sequence
	20	37.4	4.9	108759	1 CP000084_12	Continuation (13 o
	21	37.4	4.9	110000	1 CP000084_11	Continuation (12 o
	22	37.2	4.9	654	10 BV430511	BV430511 S237P6480
	23	37.2	4.9	139266	14 AC068747	AC068747 Homo sapi
	24	37	4.8	205651	14 AC128791	AC128791 Rattus no
	25	37	4.8	228747	14 AC095294	AC095294 Rattus no
	26	36.8	4.8	4791	9 BC060634	BC060634 Mus muscu
	27	36.8	4.8	196878	14 AC120071	AC120071 Rattus no
	28	36.8	4.8	198577	9 AC133189	AC133189 Mus muscu
	29	36.8	4.8	250927	14 AC154384	AC154384 Mus muscu
	30	36.6	4.8	442	10 BV184118	BV184118 sqm14314
	31	36.6	4.8	155269	14 AF191252	AF191252 Homo sapi
	32	36.6	4.8	155366	8 AC104561	AC104561 Homo sapi
	33	36.6	4.8	164132	8 AC104971	AC104971 Homo sapi
	34	36.6	4.8	173609	8 AL451077	AL451077 Human DNA
	35	36.4	4.8	223205	14 AC128730	AC128730 Rattus no
	36	36.4	4.8	230668	14 AC126903	AC126903 Rattus no
	37	36.2	4.7	10767	1 AE010844	AE010844 Methanosa
	38	36.2	4.7	110000	15 AP008207_022	Continuation (23 o
	39	36.2	4.7	128778	9 AC154750	AC154750 Mus muscu
	40	36.2	4.7	167405	15 AP002483	AP002483 Oryza sat
	41	36.2	4.7	175565	15 AP003311	AP003311 Oryza sat
	42	36	4.7	299	9 BC093588	BC093588 Mus muscu
	43	36	4.7	33281	8 AL359747	AL359747 Human DNA
	44	36	4.7	189775	9 AC112687	AC112687 Mus muscu
	45	36	4.7	207478	14 AC108599	AC108599 Rattus no

#### ALIGNMENTS

RESULT 1	OOS515523	763 bp	mRNA	linear	INV 18-APR-2003
LOCUS	OOS515523				
DEFINITION	OOS515523				
ACCESSION	AJ515523				
VERSION	AJ515523.1	GI:25005299			
KEYWORDS	aasp2 gene; ancylostoma-secreted protein-like protein.				
SOURCE	Ostertagia ostertagi				
ORGANISM	Ostertagia ostertagi				
REFERENCE	1	Geldhof, P., Vercouteren, I., Gevaert, K., Staes, A., Knox, D. P., Vandekerckhove, J., Vercruyssen, J. and Claerebout, B.			
AUTHORS	2	(bases 1 to 763)			
JOURNAL	12706806				
PUBMED	12706806				
REFERENCE	2	(bases 1 to 763)			
AUTHORS	Geldhof, P. B.				
JOURNAL	Direct Submission				
TITLE	Submitted (07-NOV-2002) Geldhof P.B., Parasitology, Ghent University, Salisburylaan 133 Merelbeke, 9820, BELGIUM				
FEATURES	location/Qualifiers				
source	1. 763				
	/organism="Ostertagia ostertagi"				
	/mol_type="mRNA"				
	/db_xref="taxon:6317"				
	/dev_stage="adult"				
	1. 763				
	/gene="aasp2"				
	11. 706				
	/gene="aasp2"				
	/codon_start=1				
	/product="ancylostoma-secreted protein-like protein"				
	/protein_id="CAD56659.1"				
	/db_xref="GI:25005300"				
	/db_xref="GO:081FT6"				
	/db_xref="UniProt/TREMBL:Q81FT6"				
	/translation="MSAAVVAVLALFSAAGFCPCNSLSQSDSARQIFLDFHNDV				



RRNIALGNGLINWTVNADAVILGPAQNMVKVMDNCNLEEVAAQOIAPCNDPLPINTSL  
AONIARMLYFKDSEETVLQOVSWYWSASLGFMKGTCLDOPANOWAEPLIANIYRN  
RKVGCAHKICPAQONMVVSCVYSPKLAIPNEVIWQEGKACVCDARPDSPCCDNLCDTR  
DAASVRHQCCASP"

ORIGIN

Query Match		100.0%;	Score 763;	DB 2;	Length 763;
Best Local Similarity		100.0%;	Pred. No. 4.5e-226;		
Matches 763;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GAGAACTGCTATGTGCGCGGCTGTTGATTGCTGTTCTCCTGCGCCCTGTTCTCTATGC	60		
DB	1	GAGAACTGCTATGTGCGCGGCTGTTGATTGCTGTTCTCCTGCGCCCTGTTCTCTATGC	60		
QY	61	CGAAGCAGGCTTTGTTGTTCCGAATAGTCTAAGCCAAAGTGAAGCGCGAGGAGATTTT	120		
DB	61	CGAAGCAGGCTTTGTTGTTCCGAATAGTCTAAGCCAAAGTGAAGCGCGAGGAGATTTT	120		
QY	121	CCTCGATTTTCACAATGATGTTCCGAAATATAGCACTTGAAATGTTGATTAAC	180		
DB	121	CCTCGATTTTCACAATGATGTTCCGAAATATAGCACTTGAAATGTTGATTAAC	180		
QY	181	GACAGTAAATGCAGACGCGGTCAATCTTGCTCAGCTCAGAACATGTACAAAGTGA	240		
DB	181	GACAGTAAATGCAGACGCGGTCAATCTTGCTCAGCTCAGAACATGTACAAAGTGA	240		
QY	241	GGATTGCACTTGGAAAGATAGACACACAGATTGCGCCATGCAATGATCCCTACC	300		
DB	241	GGATTGCACTTGGAAAGATAGACACACAGATTGCGCCATGCAATGATCCCTACC	300		
QY	301	GATTAATACCAAGCCTGGCTCAAAATATCGCTAGATGGCTGTA	360		
DB	301	GATTAATACCAAGCCTGGCTCAAAATATCGCTAGATGGCTGTA	360		
QY	361	AGAGACAGTCTGCACAAGTATCGTGTATTTGGTGAGCGCATCGCTGGATTATGA	420		
DB	361	AGAGACAGTCTGCACAAGTATCGTGTATTTGGTGAGCGCATCGCTGGATTATGA	420		
QY	421	AGGCAAGAACTTGACCAATTTGCTAACAGTGGGCTGAACCTCTAGCAAA	480		
DB	421	AGGCAAGAACTTGACCAATTTGCTAACAGTGGGCTGAACCTCTAGCAAA	480		
QY	481	CTATAGAAACCGAAAGTTGATGTGCCATAAGATCTGCCCGCTCAGCAAA	540		
DB	481	CTATAGAAACCGAAAGTTGATGTGCCATAAGATCTGCCCGCTCAGCAAA	540		
QY	541	AGTATCCTGCGGTATGGAAGCCCCAACTTGCAACGAACTTATCTGGCAGGA	600		
DB	541	AGTATCCTGCGGTATGGAAGCCCCAACTTGCAACGAACTTATCTGGCAGGA	600		
QY	601	AAAGGCTGTGTGCGACGCTCGTCAAGTCAATCTGCTGCAAACTGTGTGAC	660		
DB	601	AAAGGCTGTGTGCGACGCTCGTCAAGTCAATCTGCTGCAAACTGTGTGAC	660		
QY	661	GCGAGATGCTGCGAGTGTTCGCCACCAAGTGTGCGCGTCCCATGAAGCAAA	720		
DB	661	GCGAGATGCTGCGAGTGTTCGCCACCAAGTGTGCGCGTCCCATGAAGCAAA	720		
QY	721	GGTAGTCAACCCGAATTAATATTCATGCAAAAAA	763		
DB	721	GGTAGTCAACCCGAATTAATATTCATGCAAAAAA	763		

RESULT 2  
CQ794479 763 bp DNA linear PAT 19-APR-2004

LOCUS CQ794479  
DEFINITION Sequence 9 from Patent WO2004024769.

ACCESSION CQ794479  
VERSION CQ794479.1 GI:46407112

KEYWORDS Ostertagia ostertagi  
SOURCE Ostertagia ostertagi

ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;

REFERENCE 1  
AUTHORS Geldhof, P., Vercauteren, I., Claerebout, E., Vercruyse, J. and de Maere, V.

TITLE Ostertagia vaccine  
JOURNAL Patent: WO 2004024769-A 9 25-MAR-2004;

FEATURES  
source LOCATION/Qualifiers  
1..763

/organism="Ostertagia ostertagi"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6317"  
11..706

/note="unassigned protein product"

/codon\_start=1  
/protein\_id="CAG26292.1"

/db\_xref="GI:46407113"

/translation="MSAAVVAVLLALFSYAEAGFCCPNSLSQSDSARQIFLDFHNDV  
RRNIALGNGLINWTVNADAVILGPAQNMVKVMDNCNLEEVAAQOIAPCNDPLPINTSL  
AONIARMLYFKDSEETVLQOVSWYWSASLGFMKGTCLDOPANOWAEPLIANIYRN  
RKVGCAHKICPAQONMVVSCVYSPKLAIPNEVIWQEGKACVCDARPDSPCCDNLCDTR  
DAASVRHQCCASP"

ORIGIN

Query Match		100.0%;	Score 763;	DB 6;	Length 763;
Best Local Similarity		100.0%;	Pred. No. 4.5e-226;		
Matches 763;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GAGAACTGCTATGTGCGCGGCTGTTGATTGCTGTTCTCCTGCGCCCTGTTCTCTATGC	60		
DB	1	GAGAACTGCTATGTGCGCGGCTGTTGATTGCTGTTCTCCTGCGCCCTGTTCTCTATGC	60		
QY	61	CGAAGCAGGCTTTGTTGTTCCGAATAGTCTAAGCCAAAGTGAAGCGCGAGGAGATTTT	120		
DB	61	CGAAGCAGGCTTTGTTGTTCCGAATAGTCTAAGCCAAAGTGAAGCGCGAGGAGATTTT	120		
QY	121	CCTCGATTTTCACAATGATGTTCCGAAATATAGCACTTGAAATGTTGATTAAC	180		
DB	121	CCTCGATTTTCACAATGATGTTCCGAAATATAGCACTTGAAATGTTGATTAAC	180		
QY	181	GACAGTAAATGCAGACGCGGTCAATCTTGCTCAGCTCAGAACATGTACAAAGTGA	240		
DB	181	GACAGTAAATGCAGACGCGGTCAATCTTGCTCAGCTCAGAACATGTACAAAGTGA	240		
QY	241	GGATTGCACTTGGAAAGATAGACACACAGATTGCGCCATGCAATGATCCCTACC	300		
DB	241	GGATTGCACTTGGAAAGATAGACACACAGATTGCGCCATGCAATGATCCCTACC	300		
QY	301	GATTAATACCAAGCCTGGCTCAAAATATCGCTAGATGGCTGTA	360		
DB	301	GATTAATACCAAGCCTGGCTCAAAATATCGCTAGATGGCTGTA	360		
QY	361	AGAGACAGTCTGCACAAGTATCGTGTATTTGGTGAGCGCATCGCTGGATTATGA	420		
DB	361	AGAGACAGTCTGCACAAGTATCGTGTATTTGGTGAGCGCATCGCTGGATTATGA	420		
QY	421	AGGCAAGAACTTGACCAATTTGCTAACAGTGGGCTGAACCTCTAGCAAA	480		
DB	421	AGGCAAGAACTTGACCAATTTGCTAACAGTGGGCTGAACCTCTAGCAAA	480		
QY	481	CTATAGAAACCGAAAGTTGATGTGCCATAAGATCTGCCCGCTCAGCAAA	540		
DB	481	CTATAGAAACCGAAAGTTGATGTGCCATAAGATCTGCCCGCTCAGCAAA	540		
QY	541	AGTATCCTGCGGTATGGAAGCCCCAACTTGCAACGAACTTATCTGGCAGGA	600		
DB	541	AGTATCCTGCGGTATGGAAGCCCCAACTTGCAACGAACTTATCTGGCAGGA	600		
QY	601	AAAGGCTGTGTGCGACGCTCGTCAAGTCAATCTGCTGCAAACTGTGTGAC	660		
DB	601	AAAGGCTGTGTGCGACGCTCGTCAAGTCAATCTGCTGCAAACTGTGTGAC	660		
QY	661	GCGAGATGCTGCGAGTGTTCGCCACCAAGTGTGCGCGTCCCATGAAGCAAA	720		

Db 661 GCGAGATGCTGCGAGTGTTCGCCACCAAGTGTGGCGCTGCCCATGAAGGAAAGAAATT 720  
Qy 721 GGTAGTACCCCGAATAAATATTATGCAAAAAA 763  
Db 721 GGTAGTACCCCGAATAAATATTATGCAAAAAA 763  
RESULT 3  
CQ794471 828 bp DNA linear PAT 19-APR-2004  
LOCUS CQ794471  
DEFINITION Sequence 1 from Patent WO2004024769.  
ACCESSION CQ794471  
VERSION CQ794471.1 GI:46407104  
KEYWORDS  
SOURCE Osteragia osteragi  
ORGANISM Osteragia osteragi  
REFERENCE 1 Geldhof, P., Vercauteren, I., Claerebout, E., Vercruysse, J. and de Maere, V.  
AUTHORS Osteragia vaccine  
JOURNAL Patent: WO 2004024769-A 1 25-MAR-2004;  
FEATURES  
SOURCE location/Qualifiers  
1. 828  
/organism="Osteragia osteragi"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6317"  
11..721  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAG26291.1"  
/db\_xref="GI:46407105"  
/translation="MQALITGIALYLVLTSTNTEAGPCCPADLNTQDEARXIFLDFTN  
QVRDIAGASPLNLTGAVQMNVLGPAKMRYRMDWDCNLEAKAKAMIPCTPLPID  
TSIPQNLAQWLFQNSQEXEVLQTPTMSWVTLASRLQPDTEANITNMQIRPLSNIAN  
WQNLKVGCAHKVCKEPTGTINWVSCAYGGEVLQDNEVMDKPTCMCNAYPNPFCNN  
LCDITAAATLRKQPKST"  
ORIGIN  
Query Match 35.8%; Score 273.4; DB 6; Length 828;  
Best Local Similarity 63.6%; Pred. No. 1.9e-73;  
Matches 447; Conservative 3; Mismatches 244; Indels 9; Gaps 2;  
Qy 38 CTCCTGGCCCTGTCTCTCTATGCCGAGCAGGCTTTGTGTCCGAATAGCTAAGCCAA 97  
Db 44 CTGCTGCTGCTGACATCAATAACGAGAGGTTTGTCTGCCAGCAGATCTAAACCAA 103  
Qy 98 AGTGACAGCGCGAGGAGCAGATTTCTCGATTTCACAATGATGTCGAAATATAGCA 157  
Db 104 ACTGATGAGGCAAGAAAATCTCTCGATTTCACAATCAAGTTCGCCGTGATATAGCA 163  
Qy 158 CTGGAATGTTGTGATAACTGACAGTAATGACAGCGGTCA--TTCTGTGTCGA 214  
Db 164 GGTGCAAGCCCGTGTGCTCAACCTCACCGAGCTGTCARATGCAATGTTCTCGTCCA 223  
Qy 215 GCTCAACAATGTACAAGAGTGAGTGGATTGCACTTGGAGAGAGTAGACACACACAG 274  
Db 224 GCTAAGACATGTACAGATGAGCTGGGACTGCAATCTGGAAGCAAAAGCAAGCAATG 283  
Qy 275 ATTGCGCATGCAATGATCCCTTACCGATAAATACCAAGCTTGGCTCAAAATATGCTAGA 334  
Db 284 ATTGGCCATGCACTACGCTCTGCAATAGACAGAGTATTCACAAATCTCGCTCAR 343  
Qy 335 TGGCTGTACTTCAAGAAGACAGTGAAGAGACAGTTCTGCAACAAGTATGTTGTTATGG 394  
Db 344 TGGCTACTTTTCAAAAACAGTACAGAAAGTGAAGTGTGACGCAAAAGCCCTGTTGG 403  
Qy 395 GTGAGCGCATCGCTGGATTATGAAGAGCAGAACTTGACCAATTGCTAACCAGTGG 454  
Db 404 GTAACCGCATCACTACGAATCTTCACTGATACAGAGGTAACTTATTAATGCGCA 463

Qy 455 GCTGAACCTCTAGCAACATTCGAACATATAGAAACCGAAGTTGGATGTGCCATAAG 514  
Db 464 ATTAGACCACTATCCACATTCGCAAGTGGCAAAACCTAAAGTTGATGTCTCAGAA 523  
Qy 515 ATCTG-----CCCCGCTCAGCAAAACATGTAATCTCTCCGTTATGGAAGCCCCAA 568  
Db 524 GTGTGCAAAATTCCTCCACCGGACAAATATGTTGTGTCTTGCCTTATGCGCGCAAGTA 583  
Qy 569 CTTCACCGAAGCAAGTATCTGCGAGAGCAAGAAAGCTTGTGTGCGAGCGTCTCCA 628  
Db 584 CTCCAAAGTAAAGCAAGTGTATGGAAGAGGACCAACTTGCAATGCAATGCTTATCCC 643  
Qy 629 GATTCATCTGCTGCGACCAACCTGTGACACGCGAGATGTCGAGTGTGCGCACCAG 688  
Db 644 AACTGTTCTGCTGCAACAATCTGTGTGACACATAGCTGCTGCGACACTTCGCAAGCAG 703  
Qy 689 TGTGCGCGTCCCATGAAGCGAAAGAAATTTGTAATCACC 731  
Db 704 CCTGTAAATCGACTGAAGCGAAAGCGCTTGTGATGTCCC 746  
RESULT 4  
OOS310812 712 bp mRNA linear INV 12-MAR-2003  
LOCUS OOS310812  
DEFINITION Osteragia osteragi mRNA for ancylostoma-secreted protein-1 like  
protein (OSP-2 gene), clone F175.  
ACCESSION AJ310812  
VERSION AJ310812.2 GI:18860823  
KEYWORDS ancylostoma-secreted protein-1 like protein; OSP-2 gene.  
SOURCE Osteragia osteragi  
ORGANISM Osteragia osteragi  
REFERENCE 1 Geldhof, P., Peelaers, I., Claerebout, E., Bex, G. and Vercruysse, J.  
AUTHORS Trichostrongyloidea; Haemonchidae; Osteraglinae; Osteragia.  
TITLE Identification of excretory-secretory products of larval and adult  
Osteragia osteragi by immunoscreening of cDNA libraries  
JOURNAL Mol. Biochem. Parasitol. 126 (2), 201-208 (2003)  
PUBMED 12615319  
REFERENCE 2 Vercauteren, I.J.R.  
AUTHORS Direct Submission  
TITLE Submitted (29-MAR-2001) Vercauteren I.J.R., Department of  
Parasitology, Ghent University, Faculty of Veterinary Medicine,  
Salsburylaan, 133, B-9820 MERELBEKE, BELGIUM  
JOURNAL Revised by [3]  
REMARK 3  
AUTHORS Vercauteren, I.J.R.  
TITLE Direct Submission  
JOURNAL Submitted (13-AUG-2001) Vercauteren I.J.R., Department of  
Parasitology, Ghent University, Faculty of Veterinary Medicine,  
Salsburylaan, 133, B-9820 MERELBEKE, BELGIUM  
REMARK 4 (bases 1 to 712)  
AUTHORS Vercruysse, J.  
TITLE Direct Submission  
JOURNAL Submitted (07-FEB-2002) Vercauteren I.J.R., Department of  
Parasitology, Ghent University, Faculty of Veterinary Medicine,  
Salsburylaan, 133, B-9820 MERELBEKE, BELGIUM  
COMMENT On Feb 22, 2002 this sequence version replaced gi:13536973.  
FEATURES  
source  
1..712  
/organism="Osteragia osteragi"  
/mol\_type="mRNA"  
/db\_xref="taxon:6317"  
/clone="F175"  
/dev\_stage="adult"  
1..712  
/gene="OSP-2"  
<1..711  
/gene="OSP-2"



/codon\_start=1  
/product="ancylostoma-secreted protein-like protein"  
/protein\_id="CAD23183.1"  
/db\_xref="GI:18860824"  
/db\_xref="GOA:Q8T8F4"  
/db\_xref="UniProt/TREMBL:O8T8F4"  
/translation="MQALIGIALYLVTSNTAEGFCCPADLNQDEARXIFLDFHN  
OVRDIAGASPLNLTGAVOMRNLGPAKNMYRMDWDCNLEAKAKMIMPTPLPID  
TSIPONTAOWLFEONSQEXEVLTOTPWSWVTASLRNLQDTEANINYWQIRPLSNIAN  
WONLKVGCARHYCKEPTGTNMVVSCAVGEVLQDNEVWVDKPTCMCNAYPNSFCNN  
LCDTIAAATLRNQPKST"

ORIGIN

Query Match 34.6%; Score 263.8; DB 2; Length 712;  
Best Local Similarity 63.6%; Pred. No. 1.8e-70;  
Matches 432; Conservative 3; Mismatches 235; Indels 9; Gaps 2;

QY	38	CTCCTGGCCCTGTTCTCCTATGCGGAAGCAGGCTTTGTTGTCGGAATAGCTTAAGCCA	97
DB	34	CTGGTGTCTGTGACATCAAAATACGAAGCAGGTTTGTGCTGCCAGCAGATCTAAACCA	93
QY	98	AGTGACAGCGGAGGAGATTTCTCGATTTTCACAATGATGTCGCGAAATATAGCA	157
DB	94	ACTGATGAGGCAAGAMAAATCTTCTCGATTTTCACAATCAAGTTCGCCGTATATAGCA	153
QY	158	CTTGGAATATGTTGATTAACCTGACAGTAATGACAGCGGCTCA--TTCTTGCTCA	214
DB	154	GGTGCAAGCCCGTTGCTCAACCTCAACCGAGCTGTCARATGCGAATGTCTCGGTCA	213
QY	215	GCTCAGAACATGTACAAAGTGGAGTGGATTGCAACTTGGAGAAGTAGCAGACACAAG	274
DB	214	GCTAAGAACATGTACAGAATGACTGGGACTGCAATCTGGAAGCAAAAGCAAGGCATG	273
QY	275	ATTGGCCATGCAATGATCCCTCAACGATAATACCAAGCTGGCTCAAATATCGCTAGA	334
DB	274	ATTGGCCATGCACTACGCTCTGCGCAATAGACACAGATATCCCAAAATCTCGCTCAR	333
QY	335	TGGCTGACTTCAAGAAGAGTGAAGAGACAGTTCTGCAACAAGTATCGTGATGG	394
DB	334	TGGCTACTTTCCAAAACAGTCAAGAAATGAAGTGTGACGCAACGCCCTGCTTGG	393
QY	395	GTGAGCGCATGCGCTGGATTATGAAGGCAAGAACTTGACCAATTGCTAACCAAGTG	454
DB	394	GTAACCGCATCACTACGAATCTTCAACTGATACAGAAGCTAACATTATTAAGTGCAA	453
QY	455	GCTGAACCTCTAGCAACATTGCAAACTATAGAAACCGAAAGTTGGATGTGCCATAAG	514
DB	454	ATTAGACCACTATCCAACTTGCGAATGGCAAAACCTAAAAGTTGATGTCTCACAA	513
QY	515	ATCTG-----CCCCGCTCAGCAAAACATGTTAGTATCTGCGTGTATGGAAGCCCAA	568
DB	514	GTGTCAAAATTCGCCCAACCGGACAAATATGTTGTGTCTTGCGCTTATGGCGCGGAAGTA	573
QY	569	CTTGACCGAAGCAAGTTATCTGCAAGAGGAAAGGCTTGTGTGCGACGCTGCTCCA	628
DB	574	CTCCAAGATAAGAGTTGTATGGACAGGACCAACTTGATGCAATGCTTATCCC	633
QY	629	GATTCATTCGTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTGTGCGCACCAAG	688
DB	634	AACTCGTTCTGTGCAACAATCTGTGTGACACAATAGCTGCTGCGACACTTCGCAACCAAG	693
QY	689	TGTTGCGCGCTGCGCATGAA 707	
DB	694	CCTTGCAATCGACTTGAA 712	

RESULT 5  
AX655393 2000 bp DNA linear PAT 22-MAR-2003  
LOCUS AX655393  
DEFINITION Sequence 5263 from Patent WO03000898.  
ACCESSION AX655393  
VERSION AX655393.1 GI:29158207  
KEYWORDS

SOURCE  
ORGANISM Oryza sativa  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE  
AUTHORS 1  
Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,  
Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 5263 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES  
Source 1. 2000  
Location/Qualifiers  
1. 2000  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

ORIGIN

Query Match 5.4%; Score 41.4; DB 6; Length 2000;  
Best Local Similarity 9.3%; Pred. No. 0.4;  
Matches 54; Conservative 263; Mismatches 264; Indels 2; Gaps 1;

QY	167	GTTTGTATAACTGACAGTAATGACAGCGGCTCATCTTGTCAGCTGAGACATG	226
DB	219	RKYGTYKMTYTSASRCMARYMTTSYSWACSSYTWCRSKRSMWMMKMRKMSRSYGM	278
QY	227	TACAAAGTGACTGGATTTGCAACTTGAAGAGTAGCAGACACACAGATTGCCCATGC	286
DB	279	SWSYKMMCTAYKKSYSRWCMYRGGWGATRYWGRGYMSRMAAMYKMYWYRGYKGM	338
QY	287	AATGATCCCTACCGATAAATACCAAGCTGGCTCAAAATATCGCTAGATGCTGACTTC	346
DB	339	KRGWAGRAMMRSMCRMSKACYYMRMRMRMTRRRRMAKSSRTSRKRRKMRKRYK	398
QY	347	AAAGCAGTGAAGAGACAGTTCTGCAACAAGTATCGTGATTTGGTGAGCGCATCG	406
DB	399	RMRGYSRMRSCRARMRKCRSGRAWMGCRGCTCRMKSYGMRRKMSKRYKMS	458
QY	407	CTGGATTTATGAAGGACGAACTTGACCAATTGCTTAACCAAGTGGCTGAACCTCTA	466
DB	459	RMYRWRKKCSRTTWMGTGGMGTGRCRYKRSRKRKRRRRWRMYRMYRYM	518
QY	467	GCAACATTTGCAACTATAGAAACCGAAAGTTGG--ATGTGCCATTAAGATCTGCCCG	524
DB	519	SARYTMRVCARKVSYSAARKARCWYRGKYWAGMMKRYKRYMYKMMWYKRYKSK	578
QY	525	CTCAGCAAAACATGTTAGTATCTGCGTGTATGGAAGCCCAAACTTGACCGAAGAG	584
DB	579	SWYCKMSYYAASCMSARKAGAKCKRSKMSAWSKMSRSRKRCKRCKASKRSSAKRYAMMG	638
QY	585	TTATCTGGCAGGAAGGAAAGCTTGTGTGTGCGACGCTCGTCCAGATTCAATCTGCTGCG	644
DB	639	MTGSRMSRWKSYTCYWRKWSMKSCTTMYMYMSKYTYAKGYSWYRYRAYCMYMRWY	698
QY	645	ACAACCTGTGTGACACGCGAGATGCTGCGAGTGTTCGCCACCAAGTGTGCGCGTCCAT	704
DB	699	YRYRSYMTYMAWYTSSTRMATGMKYSGRWTSWYKYCKCSWKYRSMYTWSMWAKTW	758
QY	705	GAAAGCAAAAGAAATTGTAGTCAACCCGAATAAATATTAT 747	
DB	759	KMMRRYATRMWMMWYRYSMKWYTWCTMWGYWYMWYRYMYRMY 801	

RESULT 6  
AP008208\_049/c  
wPCOMMENT  
Sequence split into 360 fragments LOCUS AP008208 Accession AP008208  
Fragment Name Begin End  
AP008208\_000 1 11000  
AP008208\_001 100001 21000  
AP008208\_002 200001 31000  
AP008208\_003 300001 41000  
AP008208\_004 400001 51000



AP008208\_005 500001 610000  
AP008208\_006 600001 710000  
AP008208\_007 700001 810000  
AP008208\_008 800001 910000  
AP008208\_009 900001 1010000  
AP008208\_010 1000001 1110000  
AP008208\_011 1100001 1210000  
AP008208\_012 1200001 1310000  
AP008208\_013 1300001 1410000  
AP008208\_014 1400001 1510000  
AP008208\_015 1500001 1610000  
AP008208\_016 1600001 1710000  
AP008208\_017 1700001 1810000  
AP008208\_018 1800001 1910000  
AP008208\_019 1900001 2010000  
AP008208\_020 2000001 2110000  
AP008208\_021 2100001 2210000  
AP008208\_022 2200001 2310000  
AP008208\_023 2300001 2410000  
AP008208\_024 2400001 2510000  
AP008208\_025 2500001 2610000  
AP008208\_026 2600001 2710000  
AP008208\_027 2700001 2810000  
AP008208\_028 2800001 2910000  
AP008208\_029 2900001 3010000  
AP008208\_030 3000001 3110000  
AP008208\_031 3100001 3210000  
AP008208\_032 3200001 3310000  
AP008208\_033 3300001 3410000  
AP008208\_034 3400001 3510000  
AP008208\_035 3500001 3610000  
AP008208\_036 3600001 3710000  
AP008208\_037 3700001 3810000  
AP008208\_038 3800001 3910000  
AP008208\_039 3900001 4010000  
AP008208\_040 4000001 4110000  
AP008208\_041 4100001 4210000  
AP008208\_042 4200001 4310000  
AP008208\_043 4300001 4410000  
AP008208\_044 4400001 4510000  
AP008208\_045 4500001 4610000  
AP008208\_046 4600001 4710000  
AP008208\_047 4700001 4810000  
AP008208\_048 4800001 4910000  
AP008208\_049 4900001 5010000  
AP008208\_050 5000001 5110000  
AP008208\_051 5100001 5210000  
AP008208\_052 5200001 5310000  
AP008208\_053 5300001 5410000  
AP008208\_054 5400001 5510000  
AP008208\_055 5500001 5610000  
AP008208\_056 5600001 5710000  
AP008208\_057 5700001 5810000  
AP008208\_058 5800001 5910000  
AP008208\_059 5900001 6010000  
AP008208\_060 6000001 6110000  
AP008208\_061 6100001 6210000  
AP008208\_062 6200001 6310000  
AP008208\_063 6300001 6410000  
AP008208\_064 6400001 6510000  
AP008208\_065 6500001 6610000  
AP008208\_066 6600001 6710000  
AP008208\_067 6700001 6810000  
AP008208\_068 6800001 6910000  
AP008208\_069 6900001 7010000  
AP008208\_070 7000001 7110000  
AP008208\_071 7100001 7210000  
AP008208\_072 7200001 7310000  
AP008208\_073 7300001 7410000  
AP008208\_074 7400001 7510000  
AP008208\_075 7500001 7610000  
AP008208\_076 7600001 7710000  
AP008208\_077 7700001 7810000

AP008208\_078 7800001 7910000  
AP008208\_079 7900001 8010000  
AP008208\_080 8000001 8110000  
AP008208\_081 8100001 8210000  
AP008208\_082 8200001 8310000  
AP008208\_083 8300001 8410000  
AP008208\_084 8400001 8510000  
AP008208\_085 8500001 8610000  
AP008208\_086 8600001 8710000  
AP008208\_087 8700001 8810000  
AP008208\_088 8800001 8910000  
AP008208\_089 8900001 9010000  
AP008208\_090 9000001 9110000  
AP008208\_091 9100001 9210000  
AP008208\_092 9200001 9310000  
AP008208\_093 9300001 9410000  
AP008208\_094 9400001 9510000  
AP008208\_095 9500001 9610000  
AP008208\_096 9600001 9710000  
AP008208\_097 9700001 9810000  
AP008208\_098 9800001 9910000  
AP008208\_099 9900001 10010000  
AP008208\_100 10000001 10110000  
AP008208\_101 10100001 10210000  
AP008208\_102 10200001 10310000  
AP008208\_103 10300001 10410000  
AP008208\_104 10400001 10510000  
AP008208\_105 10500001 10610000  
AP008208\_106 10600001 10710000  
AP008208\_107 10700001 10810000  
AP008208\_108 10800001 10910000  
AP008208\_109 10900001 11010000  
AP008208\_110 11000001 11110000  
AP008208\_111 11100001 11210000  
AP008208\_112 11200001 11310000  
AP008208\_113 11300001 11410000  
AP008208\_114 11400001 11510000  
AP008208\_115 11500001 11610000  
AP008208\_116 11600001 11710000  
AP008208\_117 11700001 11810000  
AP008208\_118 11800001 11910000  
AP008208\_119 11900001 12010000  
AP008208\_120 12000001 12110000  
AP008208\_121 12100001 12210000  
AP008208\_122 12200001 12310000  
AP008208\_123 12300001 12410000  
AP008208\_124 12400001 12510000  
AP008208\_125 12500001 12610000  
AP008208\_126 12600001 12710000  
AP008208\_127 12700001 12810000  
AP008208\_128 12800001 12910000  
AP008208\_129 12900001 13010000  
AP008208\_130 13000001 13110000  
AP008208\_131 13100001 13210000  
AP008208\_132 13200001 13310000  
AP008208\_133 13300001 13410000  
AP008208\_134 13400001 13510000  
AP008208\_135 13500001 13610000  
AP008208\_136 13600001 13710000  
AP008208\_137 13700001 13810000  
AP008208\_138 13800001 13910000  
AP008208\_139 13900001 14010000  
AP008208\_140 14000001 14110000  
AP008208\_141 14100001 14210000  
AP008208\_142 14200001 14310000  
AP008208\_143 14300001 14410000  
AP008208\_144 14400001 14510000  
AP008208\_145 14500001 14610000  
AP008208\_146 14600001 14710000  
AP008208\_147 14700001 14810000  
AP008208\_148 14800001 14910000  
AP008208\_149 14900001 15010000  
AP008208\_150 15000001 15110000

AP008208_151	15100001	15210000
AP008208_152	15200001	15310000
AP008208_153	15300001	15410000
AP008208_154	15400001	15510000
AP008208_155	15500001	15610000
AP008208_156	15600001	15710000
AP008208_157	15700001	15810000
AP008208_158	15800001	15910000
AP008208_159	15900001	16010000
AP008208_160	16000001	16110000
AP008208_161	16100001	16210000
AP008208_162	16200001	16310000
AP008208_163	16300001	16410000
AP008208_164	16400001	16510000
AP008208_165	16500001	16610000
AP008208_166	16600001	16710000
AP008208_167	16700001	16810000
AP008208_168	16800001	16910000
AP008208_169	16900001	17010000
AP008208_170	17000001	17110000
AP008208_171	17100001	17210000
AP008208_172	17200001	17310000
AP008208_173	17300001	17410000
AP008208_174	17400001	17510000
AP008208_175	17500001	17610000
AP008208_176	17600001	17710000
AP008208_177	17700001	17810000
AP008208_178	17800001	17910000
AP008208_179	17900001	18010000
AP008208_180	18000001	18110000
AP008208_181	18100001	18210000
AP008208_182	18200001	18310000
AP008208_183	18300001	18410000
AP008208_184	18400001	18510000
AP008208_185	18500001	18610000
AP008208_186	18600001	18710000
AP008208_187	18700001	18810000
AP008208_188	18800001	18910000
AP008208_189	18900001	19010000
AP008208_190	19000001	19110000
AP008208_191	19100001	19210000
AP008208_192	19200001	19310000
AP008208_193	19300001	19410000
AP008208_194	19400001	19510000
AP008208_195	19500001	19610000
AP008208_196	19600001	19710000

Query Match 5.4%; Score 41.4; DB 15; Length 110000;  
Best Local Similarity 54.2%; Pred. No. 0.47;  
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Oy	360	AAGGACAGTCTGCAACAAGTATCGTGTAATTGGGTGAGCGCATCGCTGGATTATGA	419
Db	88566	AAGTGAAGTTCAAGATGATGCTCTTACAATCGAGTATGATGCAAGCAAAAGGA	88507
Oy	420	AAGGACGAAACTGACCAATTTGCTAACCAAGTGGGCTGAACCTCTAGCAAAACATTGCAA	479
Db	88506	TTGGAGCAGAACTAATCGATTTTACCTACAAATTGGCTGTAAATTTCACAAGCATTGCAA	88447
Oy	480	ACTATAGAACCGAAAGGTTGATGTGCCATAAG	514
Db	88446	AATGTGAATTCAACATTTCTGGCTTCAAAATATG	88412

RESULT 7  
AP005849/c 132729 bp DNA linear PLN 16-NOV-2004  
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,  
DEFINITION BAC clone:OSJNB0031B09.  
ACCESSION AP005849  
VERSION AP005849.3 GI:46390951  
KEYWORDS  
SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Erihartoideae; Oryzeae; Oryza.  
REFERENCE 1  
AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.  
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC  
clone:OSJNB0031B09  
JOURNAL Published Only in Database (2002)  
REFERENCE 2  
AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.  
TITLE Direct Submission  
JOURNAL Submitted (24-OCT-2002) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki, 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
On Apr 14, 2004 this sequence version replaced gi:34850234.  
Genes were predicted from the integrated results of the following:  
GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH  
(http://www.softberry.com/), GeneMark.hmm  
(http://opal.biology.gatech.edu/GeneMark/), Glimmer  
(http://www.tigr.org/tdb/glimmer/glmr.form.html), RiceHMM  
(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor  
(http://bioinformatics.lastate.edu/cgi-bin/sp.cgi), sjm4  
(http://globin.cse.psu.edu/html/docs/sjm4.html), gap2  
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The  
genomic sequence was searched against NCBI NonRedundant Protein  
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA  
sequence database at RGP or DDBJ. Protein homologies of the coding  
regions were searched against NCBI NonRedundant Protein database  
with BLASTP. ESTs represent the identified cDNA sequences using  
BLASTN with the corresponding DDBJ accession no. and RGP clone ID.  
Full-length cDNAs represent the identified cDNA sequences using  
BLASTN with the corresponding DDBJ accession no.  
A gene with identity or significant homology to a protein is  
classified based on the protein name to indicate the homology level  
such as same name, 'putative-' and '-like protein'. A gene without  
significant homology to any protein but with full-length cDNA or  
EST homology (covering almost the entire length of partial  
sequence) is classified as an 'unknown' protein. A gene predicted  
by two or more gene prediction programs is classified as a  
'hypothetical' protein according to IRGSP standard. A gene  
predicted by a single gene prediction program is also classified as a  
miscellaneous feature of the sequence.  
The orientation of the sequence is from -21M13 to M13rev of the BAC  
clone. This sequence of OSJNB0031B09 clone has an overlap with  
OU145\_F01 (DDBJ: AP004046) clone at 5' end and with P0453H10  
(DDBJ: AP005785) clone at 3' end. Detailed information on overlap  
and assembly quality together with annotation of this entry is  
available at http://rgp.dna.affrc.go.jp/Genomeseg.html.

FEATURES  
source  
1..132729  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="2"  
/clone="OSJNB0031B09"  
507..1259  
/gene="OSJNB0031B09.1"  
join(<507..682,797..>1259)  
/gene="OSJNB0031B09.1"  
/note="start and end point are not identified"  
join(507..682,797..1259)  
/gene="OSJNB0031B09.1"  
/note="predicted by Genemark.hmm etc."  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAD16465.1"  
/db\_xref="GI:46390952"  
/translation="MDARGARLITDGGDGDGDVGARGAERRLRGGGAALCDGAAMW

gene  
mRNA  
CDS

```

gene
misc_feature
RBRGVTBAAMWSSGVARRGARGDVTTTRRGARGGDATACGAARGARDGDTTGV
RRGARGARGRRDGRAATATATDGGGSGGARREAWSTGAATRRPGDGMARRRRW
RQARHGARAGARGSVMTAGSGGTGLGAVSARQMGVRLNVEHLAADE"
  join(2424. .2428,3888. .4800,6598. .6753,6862. .7095)
  /gene="OSJNBb0031B09.2"
  join(2424. .2428,3888. .4800,6598. .6753,6862. .7095)
  /gene="OSJNBb0031B09.2"
  /note="copla-type polypeptide-like"
  complement(7653. .8923)
  /gene="OSJNBb0031B09.3"
  complement(join(7653. .7929,8479. .8923))
  /gene="OSJNBb0031B09.3"
  /note="contains EST(s): AU161887(S6190),D39569(S1037)
  contains full-length cDNA(s): AK120425"
  /codon_start=1
  /product="ribosomal protein s21-like protein"
  /protein_id="BAD16466.1"
  /db_xref="GI:46390953"
  /translation="MAAPATPILLSALLQLPLAFSGRRSPPEPSVHVAPRSPTA
  AVATKGYNVQIVDENEGEESIERRFRREVRAVLQEIKRARRYESKDERKRRAR
  EAGRRNRRRMEERPFPEDDADSEARRARDEKONWEIDGIL"
  complement(9908. .13128)
  /gene="OSJNBb0031B09.4"
  complement(join(9908. .10162,11120. .11325,12777. .12893,
  13018. .13128))
  /gene="OSJNBb0031B09.4"
  /note="supported by full-length cDNA(s): AK111066"
  complement(join(11149. .11325,12777. .12893,13018. .13071))
  /gene="OSJNBb0031B09.4"
  /note="contains EST(s): AU092625(C52617),AU063035(C52617)
  contains full-length cDNA(s): AK111066
  similar to Arabidopsis thaliana chromosome 3, At3g01435"
  /codon_start=1
  /product="unknown protein"
  /protein_id="BAD16467.1"
  /db_xref="GI:46390954"
  /translation="MLPQSQSSSLQRLNHVEQMTVRAVNLACTVMEELGNATGPRTEG
  VAGHCRFELAMKEIQTLREBIKISACEYRPFKCDYSARIANEICCKVEYVLEKLD
  AMQNIIECSTGS"
  13885. .15378
  /gene="OSJNBb0031B09.5"
  join(13885. .14326,14395. .15378)
  /gene="OSJNBb0031B09.5"
  /note="supported by full-length cDNA(s): AK108567"
  join(13898. .14326,14395. .15099)
  /gene="OSJNBb0031B09.5"
  /note="contains full-length cDNA(s): AK108567"
  /codon_start=1
  /product="putative Anter-specific proline-rich protein APG
  precursor"
  /protein_id="BAD16468.1"
  /db_xref="GI:46390955"
  /translation="MRSRFDLVAVALLQLSVLLLSGEPAAAAGSKVAATIVFGD
  STDVTGNMYLSTLVRSDFAPYGRDLQLAGAGVSGGNGRPTGRFSNGRLAVDFISEA
  FGLPPLVPAYLDPAVNMSSLAGACFASAGAGYDNATSDLSVLPLMKELDYFKEYAA
  RLRSFRGDDDDAAAAAATLSEALYIVSMGTNDFLENTYAVARGHAAEYSTAAAYGDY
  LIGVAAFAVRELHAIAGARYDNLGPEMGLPLBRATGCACTEEYNAAVAGRFNAGLQD
  MIARLNGELGGARIYGDVYGAVAAVLADPAAYGVENYKAGCCGVTGVPEMGYMGCA
  GARSPLTCTDASGKAFMDAIHPTERLHRAIADAKMNTLHVL"
  15811. .19509
  /gene="OSJNBb0031B09.6"
  join(15811. .16277,17884. .18385,19169. .19509)
  /gene="OSJNBb0031B09.6"
  /note="supported by full-length cDNA(s): AK060691"
  join(15882. .16277,17884. .18385,19169. .19365)
  /gene="OSJNBb0031B09.6"
  /note="contains full-length cDNA(s): AK060691"
  /codon_start=1
  /product="putative Anter-specific proline-rich protein APG
  precursor"

```

```

/protein_id="BAD16469.1"
/db_xref="GI:46390956"
/translation="MASPLVRLLLLLLVVAASRGAAAKAVRTAVIVFGDST
VDGNNNOIGTPLRSDPPPYGRDMGGARATGRFNGRLAPDFMSESLGPLVPAYL
DPAYGIADPARGVCFPASAGTGLDNATAGLSVIPLWKEVEYREYORLRHAGAAA
RDVVRGALHVVISGTNDPLENTFLNYLATGRFARYSVGEYEDYLVAARAFLAIHRLGA
RVTFPAGSPMGCLPLERTAGALGGGGCVBEYNRVAREYNGYEAMVRSIABLPL
RLKVAFIPYDNMDLITHEPKYGLNEVEEGCCATGRPFEMGFMCNDESPLTCDASKY
LPWFDAFPTEKVNRIQAHTLDVCQQGVLL"
complement(20777..23474)
/gene="OSJNB0031B09.7"
complement(join(20777..20936,21084..21920,23153..23474))
/gene="OSJNB0031B09.7"
/note="supported by full-length cDNA(s): AK108100"
complement(join(20901..20936,21084..21920,23153..23158))
/gene="OSJNB0031B09.7"
/note="contains EST(s): AU183867(B51076),AU183866(B51076)
contains full-length cDNA(s): AK108100"
/codon_start=1
/product="unknown protein"
/protein_id="BAD16470.1"
/db_xref="GI:46390957"
/translation="MAGQGPVPRKGSWDENIGYTVFPDKYRENKTAPTPATAPAT
HDDYEPDYEHYENLSRKASRPSSHGHHGHPAPAHHHHRAPPPOOOHH
HYPASQRSGNGYHRSGSNGSSAASEASSRGSFKSPKPYPQRYSSNSNSGGGGYG
YCGPGAGAYAAPPMQHMHQOQHAQPRVAASPRLHALPOVSRAKAASAVPKFGVM
DEQNAAAAGQGFTVOFEKVKRRHREVAKAAAAAGPDVTPRMSPAQAAPRHPRKABKR
SFPSKYRCMFPRVRE"
complement(25261..26644)
/gene="OSJNB0031B09.8"
complement(join(<25261..25446,26001..26129,26630..>26644))
/gene="OSJNB0031B09.8"
/note="start and end point are not identified"
complement(join(25261..25446,26001..26129,26630..26644))
/gene="OSJNB0031B09.8"
/note="predicted by FGENESH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD16471.1"
/db_xref="GI:46390958"
/translation="MGKDXYLVVALVPTDPATDYLVPYVVKCMSRLMGNKYNDLF
GYFPLLQLIRVTACSPARRSKASHVPDVAPVVCNSGDNDGDGYNNKSWALVMKLISV
IGIRKLFP"
29116..29268
/gene="OSJNB0031B09.9"
29116..29268
/misc_feature
/gene="OSJNB0031B09.9"

Query Match      5.4%; Score 41.4; DB 15; Length 132729;
Best Local Similarity 54.2%; Pred. No. 0.47;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

gene
29116..29268
/gene="OSJNB0031B09.9"
29116..29268
/misc_feature
/gene="OSJNB0031B09.9"

Oy    360 AAGAGACAGTTCTGCACAAGTATCGTGTATTGGGTGAGCGCATCGCTGGATTATGA 419
      ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db    59692 AAGTGAAGTCAGAATGATGCTCCTTACAATCGAGTAGCATCGCAAGCAAAGGA 59633
      ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Oy    420 AAGCACGAACCTTGACCAATTGCTAACCAAGTGGGCTGAACCTTACGAACAATTGCCAA 479
      ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db    59632 TTGGAGCAGAACTTAATCGATTTCCTTACAATTGGCTTGTAATTTCACAAAGCATTGCCAA 59573
      ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Oy    480 ACTATAGAAACCGAAAGGTTGGATGTGCCATAAG 514
      ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db    59572 AATGTGAATTCACACTTCTCGGCTTCCAATATATG 59538
      ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
LOCUS       AC131303               187110 bp     DNA          linear   ROD 02-SEP-2004
DEFINITION Mus musculus chromosome 1, clone RP24-119019, complete sequence.
ACCESSION   AC131303
VERSION     AC131303.8   GI:51854578
KEYWORDS
SOURCE      Mus musculus (house mouse)

```



ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

1 (bases 1 to 187110)  
Birren,B., Nusbaum,C. and Lander,B.  
Mus musculus chromosome 1, clone RP24-119019  
Unpublished  
2 (bases 1 to 187110)  
Birren,B., Nusbaum,C., Lander,B., All,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 187110)  
Birren,B., Nusbaum,C., Lander,B., Abouelleil,A., Allen,N.,  
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,  
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,  
Dearellano,K., Diaz,J.S., Dodge,S., Doooley,K., Dorris,L.,  
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,  
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N.,  
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
Zimmer,A. and Zody,M.

Direct Submission  
Submitted (10-AUG-2004) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 187110)  
Birren,B., Nusbaum,C., Lander,B., Abouelleil,A., Allen,N.,  
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,  
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,  
Dearellano,K., Diaz,J.S., Dodge,S., Doooley,K., Dorris,L.,  
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,  
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N.,

TITLE  
JOURNAL

COMMENT

Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (02-SEP-2004) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 2, 2004 this sequence version RepeatMasker replaced  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@broad.mit.edu](mailto:sequence_submissions@broad.mit.edu)  
----- Project Information  
Center project name: L27079  
Center clone name: 119\_O\_19  
-----

Some of the sequence contained within base pairs 1 - 121827 was  
stolen from accession AC123556.

FEATURES

source

1. 187110  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"  
/clone="RP24-119019"  
/clone\_lib="RPCI-24 Male Mouse BAC"  
1. 4  
/note="clone boundary  
clone\_end:SP6  
site:MboI"  
241. 272  
/rpt\_family="(TG)n"  
385. 442  
/rpt\_family="(TG)n"  
688. 713  
/rpt\_family="(TTTTCG)n"  
1016. 1251  
/rpt\_family="URRIA"  
1932. 2018  
/rpt\_family="(TG)n"  
2348. 2378  
/rpt\_family="AT\_rich"  
2844. 3688  
/rpt\_family="Lx"  
3680. 4391  
/rpt\_family="Lx"  
/rpt\_family="Lx"  
4395. 4519  
/rpt\_family="(GAAA)n"  
4606. 5194  
/rpt\_family="Lx8"  
5609. 5923  
/rpt\_family="Lx8"  
7795. 7881  
/rpt\_family="(TA)n"  
7889. 7938  
/rpt\_family="(CA)n"  
8324. 8420  
/rpt\_family="B3"  
8421. 8504  
/rpt\_family="B4"  
9357. 9752  
/rpt\_family="L1\_MM"  
10446. 10649  
/rpt\_family="B3"  
11004. 11123  
/rpt\_family="MER77"  
complement(11381. 11455)

```
repeat_region /rpt family="MTD" complement(11921. .12063)
repeat_region /rpt family="B1_MM" complement(15578.16084)
repeat_region /rpt family="SLTR11A" 16239. .16403
repeat_region /rpt family="RSINB1" 17073. .17094
repeat_region /rpt family="AT_r1ch" 17100. .17249
repeat_region /rpt family="B1_MM" 17504. .17612
repeat_region /rpt family="(TTTC)n" 17637. .17657
repeat_region /rpt family="AT_r1ch" 17692. .18265
repeat_region /rpt family="L1_MM" 18337. .18363
repeat_region /rpt family="AT_r1ch" complement(18812. .21567)
repeat_region /rpt family="L1VL2" 21578. .21710
repeat_region /rpt family="(TTTC)n" complement(21745. .21880)
repeat_region /rpt family="B1_MM" 21886. .22170
repeat_region /rpt family="MTD" complement(22689. .22899)
repeat_region /rpt family="B3" 22955. .23104
repeat_region /rpt family="B1_MM" 23105. .23124
repeat_region /rpt family="(A)n" 23523. .23607
repeat_region /rpt family="MLT2B3" 23692. .24043
repeat_region /rpt family="MLT2B3" 24990. .25418
repeat_region /rpt family="L1_MM" 25443. .25704
repeat_region /rpt family="L1_MM" 26072. .26146
repeat_region /rpt family="(TCTCTG)n" complement(26204. .26233)
repeat_region /rpt family="B3A" 26234. .26267
repeat_region /rpt family="(TA)n" complement(26268. .26415)
```

Query Match	5.4%	Score 41.2;	DB 9;	Length 187110;
Best Local Similarity	49.5%;	Pred. No. 0.56;		
Matches 106;	Conservative	0;	Mismatches 108;	Indels 0;
			Gaps	0;

Oy	48	TGTTCTCCTATGCCGAAGCAGGCTTTGTTCGGATAGTCTAAGCCAAAGTGACAGCG	107
Db	73210	TTAATTCAACAGTGTGTAAAAAGGCTAATTATCTGTGAAAAGCTGACTCAAGTAACAGAA	73269
Oy	108	CGAGGCAGATTTTCTCGATTTTCACAATGATGTTCTCGAAATAAGCACTTGGAAATG	167
Db	73270	TTACAGAATGTTTCATGCAATTTTATTTTAATAAAACTCAAAGCATTTACTTCAGCATC	73329
Oy	168	GTTTGATAAACCTGCACGTAATGCAGACGGCGTCATTCTTTGGTCAGCTCAGAACATGT	227
Db	73330	CTAAGGAATCTGCATACTCAAGGCACGAGGAAGACATTTGAAACTTAAACAAAATTA	73389
Oy	228	ACAAAGTGACTGGGATTTGCCAATTGGAAAGT	261
Db	73390	ACTCTGTAGGGGTGTGATAGTGTCTTGACATTAAT	73423

RESULT 9	
AC123556/c	
LOCUS	AC123556
	198900 bp
	DNA
	linear
	ROD 08-NOV-2003

DEFINITION	Mus musculus BAC clone RP23-176M9 from 1, complete sequence
ACCESSION	AC123556
VERSION	AC123556.4
KEYWORDS	GI:29244829
SOURCE	HTG.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

REFERENCE AUTHORS TITLE
1 (bases 1 to 198900) Dauphin, S., Meyer, R., Kohlberg, S. and Haakenson, W. The sequence of <i>Mus musculus</i> BAC clone RP23-176M9

JOURNAL	Unpublished (2001)
REFERENCE	2 (bases 1 to 198900)
AUTHORS	Wilson, R.

TITLE	Sequencing of <i>Mus musculus</i>
JOURNAL	Unpublished (2001)
REFERENCE	3 (bases 1 to 198900)

**AUTHORS** McPherson, J.D. and Waterston, R.H.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (30-MAY-2002) Genome Sequencing Center. 4444 Forest Park

**REFERENCE**  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 198900)  
McPherson J D and Waterson P H

NOTES  
MELTZER, D. C. and MELTZER, K. H.  
TITLE  
Direct Submission  
Submitted (11-FEB-2003) Genome Sequencing Center, 4444 Forest Park  
Parkway St Louis MO 63108 USA

REFERENCE  
AUTHORS  
McPherson, J.D. and Waterston, R.H.  
5 (bases 1 to 198900)  
fairway, St. Louis, MO 63106, USA

**TITLE** Direct Submission  
**JOURNAL** Submitted (26-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE	6 (bases 1 to 198900)
AUTHORS	Wilson, R.
TITLE	Direct Submission

**JOURNAL** Submitted (08-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, US

**COMMENT** On Mar 26, 2003 this sequence version replaced qi:28302047.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC

Web site: <http://genome.wustl.edu>  
Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)

---

Summary Statistics

Center project name: M\_BA0176M09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see

**SOURCE INFORMATION:**  
<http://genome.wustl.edu>

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong  
(<http://www.chori.org>) from female C57BL/6J mouse kidneys and/or

brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or [Robert.A.Tordt@researchgenetics.com](mailto:Robert.A.Tordt@researchgenetics.com).

NEIGHBORING SEQUENCE INFORMATION:  
This sequence is the entire insert of the clone. This clone is overlapped by AC117187.

FEATURES

Source

Location/Qualifiers

1..198900

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="1"

/map="1"

/clone="RP23-176M9"

/clone\_1lb="RPCT-23"

2138..2200

/rpt\_family="ERV1"

3436..3494

/rpt\_family="ERV1"

4075..4320

/rpt\_family="L1"

4478..4549

/rpt\_family="ERV1"

4555..4803

/rpt\_family="L1"

4883..4937

/rpt\_family="ERV1"

4950..5305

/rpt\_family="L1"

5315..5367

/rpt\_family="ERV1"

5620..5822

/rpt\_family="L1"

6086..6363

/rpt\_family="ERVK"

6196..6437

/rpt\_family="ERVK"

6336..6616

/rpt\_family="ERVK"

6641..6691

/rpt\_family="L1"

7065..7180

/rpt\_family="B4"

7816..7956

/rpt\_family="B4"

12409..12486

/rpt\_family="Alu"

13107..13275

/rpt\_family="B2"

13518..13601

/rpt\_family="B4"

13532..13706

/rpt\_family="B2"

13743..13908

/rpt\_family="L1"

14060..14166

/rpt\_family="L1"

14150..15342

/rpt\_family="L1"

15729..16048

/rpt\_family="MALR"

17335..17449

/rpt\_family="Alu"

18936..19035

/rpt\_family="L1"

19041..19287

/rpt\_family="B4"

23075..23454

/rpt\_family="MALR"

23539..23935

/rpt\_family="L1"

24005..24090

/rpt\_family="L1"

24219..24279

/rpt\_family="L1"

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

25080..25179

/rpt\_family="B4"

repeat\_region

25485..25580

/rpt\_family="L1"

repeat\_region

25888..25959

/rpt\_family="L1"

repeat\_region

27683..27769

/rpt\_family="L1"

repeat\_region

27809..27957

/rpt\_family="L1"

repeat\_region

28112..28207

/rpt\_family="Alu"

repeat\_region

28720..29059

/rpt\_family="L1"

repeat\_region

29081..29175

/rpt\_family="Alu"

repeat\_region

30059..30195

/rpt\_family="L1"

repeat\_region

30255..30415

/rpt\_family="B2"

repeat\_region

30791..31062

/rpt\_family="L1"

repeat\_region

31275..31462

/rpt\_family="B2"

repeat\_region

31503..31701

/rpt\_family="L1"

repeat\_region

31751..31880

/rpt\_family="L1"

repeat\_region

32178..32531

/rpt\_family="L1"

repeat\_region

32667..32856

/rpt\_family="L1"

repeat\_region

32838..32933

/rpt\_family="L1"

repeat\_region

33219..33519

/rpt\_family="L1"

repeat\_region

33804..33873

/rpt\_family="L1"

repeat\_region

33875..34894

/rpt\_family="L1"

repeat\_region

34940..35077

/rpt\_family="B2"

repeat\_region

37436..37719

/rpt\_family="B4"

repeat\_region

38077..38225

/rpt\_family="B4"

repeat\_region

38299..38443

/rpt\_family="MALR"

repeat\_region

38655..38980

/rpt\_family="MALR"

repeat\_region

40754..40910

/rpt\_family="L1"

repeat\_region

41010..41141

repeat\_region

Query Match 5.4%; Score 41.2; DB 9; length 198900;

Best Local Similarity 49.5%; Pred.No.0.56;

Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 48 TGTTCCTATGCGGAGCAGGCTTTGTTGTCGGAATAGTCTAAGCCAAAGTGACAGCG 107

DB 48619 TGATTTCACAGTGTGTAAAGGCTAATTATCTGTGAAAGTCTGACCTCAAGTAACAGAA 48560

QY 108 CGAGCAGATTTTCTCGATTTTCAATGATGTTGCGAATAATAGCACTTGGAAATG 167

DB 48559 TTACAGAAATGTCATGATTTTAAATAAATACTCAAGAATTAAGCACTTCAAGCATC 48500

QY 168 GTTGTAACTGACAGTAATGACAGCGGTCATTTCTGTCAGACTCAGAACATGT 227

DB 48499 CTAAAGAAATCTGCATAGTCAAGCAGAGGAAGACATTTGAACCTTAAACAAAATAA 48440

QY 228 ACAAGTGACTGGATTCGAAGTCTTGAAGT 261

DB 48439 ACTCTGAGGGTGTAGTAGTCTTGACATTAAT 48406





[illegible]

LOCUS	166494		7218 bp	DNA	linear	PAT 28-DEC-1997
DEFINITION	Sequence 14 from patent US 5670367.					
ACCESSION	166494					
VERSION	166494.1		GI:2724471			
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 7218)					
AUTHORS	Dorner, F., Schefflinger, F. and Falkner, F. Gunter.					
TITLE	Recombinant fowlpox virus					
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;					
FEATURES	Location/Qualifiers					
Source	1..7218 /organism="unknown" /mol_type="unassigned DNA"					
ORIGIN						
Query Match	5.1%;	Score 39.2;	DB 6;	Length 7218;		
Best Local Similarity	7.7%;	Pred. No. 2;				
Matches	32;	Conservative 197;	Mismatches 185;	Indels 0;	Gaps 0;	
Oy	102	ACAGCGGAGCAGATTTCCTCGATTTCACAATGATGTTGTCGAATATAGCACTTG	161			
Db	1505	AAACGGCATGTAGGCATCACTGTAATTACTATCTATGCAAGTAGTTAAGAGATAGAA	1446			
Oy	162	GAATGTTTGAATACTGACAGTAATGACAGCGGTCATTCTTGTCAGCTCAGA	221			
Db	1445	GAATTTGGTACRRR	1386			
Oy	222	ACATGTACAAAGTGCAGCTGGGATTGCCACTTGGAAGAGTAGCACACACAGATTGGC	281			
Db	1385	RRR	1326			
Oy	282	CATGCAATGATCCCTAACCGATAATACCAGCCTGGCTCAAATATCGCTAGATGGCTGT	341			
Db	1325	RRR	1266			
Oy	342	ACTTCAAGACAGTGAAGAGACAGTTCGCAAGTAGTCTGTATTGGTGAGCG	401			
Db	1265	RRR	1206			
Oy	402	CATCGCTGGGATTATGAAGGACGAACTTGACCAATTGCTAACCACTGGGCTGAC	461			
Db	1205	RRR	1146			
Oy	462	CTTAGCAAAACATTGCAAACTATAGAAACGAAAGGTGATGTCCCATAAGA	515			
Db	1145	RRR	1092			
RESULT 12						
LOCUS	AC022202	95097 bp	DNA	linear	PRI 08-OCT-2003	
DEFINITION	Homo sapiens BAC clone RP11-321B8 from 7, complete sequence.					
ACCESSION	AC022202					
VERSION	AC022202.12	GI:13992747				
KEYWORDS	HTG.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Bukacynska, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheraia; Euarchontoglires; Primates; Catarrhini;					
TITLE	Homidae; Homo.					
JOURNAL	1 (bases 1 to 95097)					
PUBMED	Suleston, J.E. and Wilson, R.					
REFERENCE	Toward a complete human genome sequence					
AUTHORS	Genome Res. 8 (11), 1097-1108 (1998)					
TITLE	9847074					
JOURNAL	2 (bases 1 to 95097)					
PUBMED	Abbott, A., Hawkins, M. and Meyer, R.					
REFERENCE	The sequence of Homo sapiens BAC clone RP11-321B8					
AUTHORS	Unpublished (2001)					

REFERENCE 3 (bases 1 to 95097)  
AUTHORS Waterston, R. H.  
TITLE Direct Submission  
JOURNAL Submitted (26-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 4 (bases 1 to 95097)  
AUTHORS Waterston, R. H.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAY-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 5 (bases 1 to 95097)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE 6 (bases 1 to 95097)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (29-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE 7 (bases 1 to 95097)  
AUTHORS Wilson, R.  
TITLE Direct Submission  
JOURNAL Submitted (08-OCT-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT On May 8, 2001 this sequence version replaced gl:13431144.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: saplens@watson.wustl.edu  
----- Summary Statistics  
Center project name: H\_NH0321E08  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see  
http://www.nhgri.nih.gov/DIR/CTB/CHR7, send  
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu

SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org  
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-468B6, 200 bp overlap the clone sequenced to the right is RP11-561N12. Actual start of this

FEATURES	clone is at base position 61428 of RP11-468B6 actual end is at base position 95097 of RP11-321E8.
source	location/Qualifiers
1. 95097	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7"
	/clone="RP11-321E8"
	/clone_lib="RPCI-11"
repeat_region	273..564
	/rpt_family="Alu"
repeat_region	1280..1352
	/rpt_family="L2"
repeat_region	1371..1404
	/rpt_family="MER1_type"
repeat_region	1570..1808
	/rpt_family="L2"
repeat_region	1861..2172
	/rpt_family="Alu"
repeat_region	3218..3294
	/rpt_family="CR1"
repeat_region	3320..3610
	/rpt_family="Alu"
repeat_region	4987..5293
	/rpt_family="Alu"
repeat_region	5727..6021
	/rpt_family="Alu"
repeat_region	6134..6454
	/rpt_family="Alu"
repeat_region	6798..6945
	/rpt_family="MIR"
repeat_region	9181..9201
	/rpt_family="AT_rich"
repeat_region	9242..9471
	/rpt_family="L1"
repeat_region	9472..9785
	/rpt_family="Alu"
repeat_region	9786..9898
	/rpt_family="L1"
repeat_region	9907..10485
	/rpt_family="ERV1"
repeat_region	10594..10760
	/rpt_family="L1"
repeat_region	11288..11311
	/rpt_family="(TAAA)n"
repeat_region	13027..13325
	/rpt_family="Alu"
repeat_region	14024..14329
	/rpt_family="Alu"
repeat_region	14395..14416
	/rpt_family="(CAA)n"
repeat_region	14818..14837
	/rpt_family="(GAAA)n"
repeat_region	15486..15770
	/rpt_family="Alu"
repeat_region	16662..16946
	/rpt_family="L1"
repeat_region	18463..18615
	/rpt_family="MER1_type"
repeat_region	18609..18650
	/rpt_family="MIR"
repeat_region	18773..18987
	/rpt_family="L1"
repeat_region	19997..20162
	/rpt_family="MER1_type"
repeat_region	20282..20582
	/rpt_family="Alu"
repeat_region	20586..20743
	/rpt_family="L1"
repeat_region	20770..21530
	/rpt_family="L1"



repeat\_region 21732..21798  
/rpt\_family="L1"  
repeat\_region 21799..22104  
/rpt\_family="Alu"  
repeat\_region 22105..22878  
/rpt\_family="L1"  
repeat\_region 22879..22902  
/rpt\_family="AT\_rich"  
repeat\_region 22934..22957  
/rpt\_family="AT\_rich"  
repeat\_region 23000..23270  
/rpt\_family="Alu"  
repeat\_region 23481..23523  
/rpt\_family="AT\_rich"  
repeat\_region 23738..23769  
/rpt\_family="AT\_rich"  
repeat\_region 24080..24144  
/rpt\_family="L1"  
repeat\_region 24145..24452  
/rpt\_family="Alu"  
repeat\_region 24453..25442  
/rpt\_family="L1"  
repeat\_region 26505..26544  
/rpt\_family="(TTTG)n"  
repeat\_region 27297..27572  
/rpt\_family="L1"  
repeat\_region 27573..27871  
/rpt\_family="Alu"  
repeat\_region 27872..28013  
/rpt\_family="L1"  
repeat\_region 28014..28325  
/rpt\_family="Alu"  
repeat\_region 28326..28773  
/rpt\_family="L1"  
repeat\_region 28777..28812

Query Match 5.1%; Score 39; DB 8; Length 95097;

Best Local Similarity 53.6%; Pred. No. 2.6;  
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 82 GAATAGTCTAAGCCAAAGTGACAGCGGAGGCGAGATTTCCTCGATTTCACAATGATGT 141  
Db 11138 GCAAGTTTATGACACTTTCMAAAGAACACATATTTTCCAGTAITTTCAACAAAAAAT 11197  
Qy 142 TCGTGAATATATAGCACTTGGAATGTTTGATTAACGTGACAGTAATGCAGACGGCGT 201  
Db 11198 TCATCGAATGTCAGACTTTAAAAAGACTTTTAAAAAATAGTTTAAAGAGCTCAAGA 11257  
Qy 202 CATCTTGTCCAGCTCAGACATGTACAA 232  
Db 11258 GATGCAAGAAATCTGAAACCAATACAA 11288

RESULT 13  
AC121380 2/c

Sequence split into 4 fragments LOCUS AC121380 Accession AC121380

Fragment Name Begin End  
AC121380\_0 1 110000  
AC121380\_1 100001 210000  
AC121380\_2 200001 310000  
AC121380\_3 300001 407488  
Continuation (3 of 4) of AC121380 from base 200001 (AC121380 Rattus norvegicus clone CH2

Query Match 5.1%; Score 38.6; DB 14; Length 110000;  
Best Local Similarity 47.7%; Pred. No. 3.5;  
Matches 113; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 126 ATTTCAATGATGTTGTCGAATATAGCACTTGAATGTTTGATAAATGACAG 185  
Db 30506 AGTATCAAAAGCATTTATAAGTAATATGCAATATAAAATTAATTAATGATAA 30447  
Qy 186 TAAATGACAGCGGTCATTTCTGTCCAGCTCAGAACATGTACAAAGTGGGATT 245

Db 30446 TAATTATTATGGGACTTTTGTGACTTAGCTAAATAATTACAAATTTTACCAAAATT 30387  
Qy 246 GCAACTTGGAGAAGTAGACACACAGATTGCGCATGCAATGATCCCTACCGATAA 305  
Db 30386 ATAAATTTTACCAATTATTCAAATTGTAGAATAAGAAATAAATGTTCACTTAAAGAAATT 30327  
Qy 306 ATACCAGCTGGCTCAAAATATCGCTAGATGCTGTACTTCAAGAAGCTGAAGAAG 362  
Db 30326 TCAGTATCCAAAGGGAATTTGAAATGTCTAGCTTAAACAGAAATTCCTAAATCAG 30270

RESULT 14  
AC120958/c  
LOCUS  
DEFINITION  
AC120958 225475 bp DNA linear HTG 19-NOV-2002  
Rattus norvegicus clone CH230-25B23, \*\*\* SEQUENCING IN PROGRESS

AC120958 GI:25090660  
HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ENRICHED.  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 225475)  
Muzny,D.Marie, Metzker,M.Lee, Abramson,S., Adams,C., Alder,J.,  
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,  
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,B., Baden,H.,  
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,  
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,  
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,  
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,  
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
Fraser,C.M., Gabriel,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,  
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,  
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,  
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,  
Hollins,B., Howell,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,  
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,  
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,  
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,  
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,  
Plopper,F., Poindexter,A., Popovic,D., Primus,B., Pu,L.-L.,  
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,  
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,  
Shetty,J., Shvartsbeyn,A., Slason,I., Sitter,C.D., Smajs,D.,  
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,  
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,  
Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,  
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
Williams,G., Willison,R., Wleczyk,R., Wooden,H., Worley,K.,  
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von







GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2006, 07:02:49 ; Search time 589 Seconds  
(without alignments)  
8633.549 Million cell updates/sec

Title: US-10-527-771-9

Perfect score: 763  
Sequence: 1 gagaactgctatgtcgcg...tcatgcaaaaaaaaaa 763

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_21: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*  
14: geneseqn2005s: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	763	100.0	763	12	ADK13846	Adk13846 Osterlagi
2	273.4	35.8	828	12	ADK13844	Adk13844 Osterlagi
3	41.4	5.4	2000	8	ADA71938	Ada71938 Rice gene
4	37.4	4.9	2000	8	ADA71938	Ada71938 Rice gene
5	36	4.7	260803	13	ABD32730	Abd32730 Human can
6	35.4	4.6	1537	6	ABN96867	Abn96867 Human glu
7	35.4	4.6	7769	6	ABN96867	Abn96867 Gene #336
8	34.6	4.5	110000	13	ABD32627_0	Abd32627 Mouse can
9	34.6	4.5	194883	11	ACN44398	Acn44398 Human gen
10	34.4	4.5	1003	6	ABQ24108	Abq24108 Oligonuc1
11	34.4	4.5	1003	6	ABQ24109	Abq24109 Oligonuc1
12	34.2	4.5	110000	13	ABD32806_1	Abd32806_1 Contnuation (2 of
13	34	4.5	3457	4	ABL10900	Ab110900 Drosophi1
14	34	4.5	4092	4	ABL18190	Ab118190 Drosophi1
15	34	4.5	7160	4	ABL13492	Ab113492 Drosophi1
16	33.6	4.4	2000	6	ABZ16370	Abz16370 Arabidops
17	33.6	4.4	130263	6	ABK83573	Abk83573 Human CDN
18	33.4	4.4	1321	14	ADY37283	Ady37283 Hookworm
19	33.4	4.4	11204	4	AA104557	Aa104557 Human rep

20	33.4	4.4	11204	4	ABL97480	Ab197480 Human tes
21	33.4	4.4	19802	4	AAK90894	Aak90894 Human dig
22	33.4	4.4	19802	4	AA104554	Aa104554 Human rep
23	33.4	4.4	19802	4	ABL97477	Ab197477 Human tes
24	33.4	4.4	21403	4	AAK90893	Aak90893 Human dig
25	33.4	4.4	21403	4	AA104555	Aa104555 Human rep
26	33.4	4.4	21403	4	ABL97478	Ab197478 Human tes
27	33.4	4.4	23763	4	AAK90895	Aak90895 Human dig
28	33.4	4.4	23763	4	AA104558	Aa104558 Human rep
29	33.4	4.4	23763	4	ABL97481	Ab197481 Human tes
30	33.4	4.4	67403	13	ADS36460	Ads36460 Human aut
31	33.4	4.4	97662	3	AAF83908	Aaf83908 Genomic s
32	33.4	4.4	106746	4	AAA10225	Aa10225 Human PCT
33	33.2	4.4	530	8	ABX98607	Abx98607 Rice leaf
34	33	4.3	804	12	ADN72288	Adn72288 Thale cre
35	33	4.3	1193	13	ADX31759	Adx31759 Plant ful
36	33	4.3	1510	3	AAC42522	Aac42522 Arabidops
37	33	4.3	1884	11	ADN39685	Adn39685 Cancer/an
38	33	4.3	2876	11	ADN39684	Adn39684 Cancer/an
39	33	4.3	44978	13	ADS36458	Ads36458 Human aut
40	32.8	4.3	3100	6	AAD22128	Aad22128 Arabidops
41	32.6	4.3	796	10	ADK58398	Adk58398 Plant DNA
42	32.6	4.3	796	10	ADK54354	Adk54354 Plant DNA
43	32.6	4.3	801	10	ADK57570	Adk57570 Plant DNA
44	32.6	4.3	1278	6	AAD33756	Aad33756 Mouse Mrg
45	32.6	4.3	1278	12	ADH08544	Adh08544 DNA seque

#### ALIGNMENTS

RESULT 1	ADK13846	ADK13846 standard; cDNA; 763 BP.
ID	ADK13846;	
AC	ADK13846;	
XX		
DT	03-JUN-2004	(first entry)
XX		
DE	Ostertagia ostertagi immunogenic protein cDNA seq id 9.	
XX		
KW	antiparasitic; gene therapy; vaccine; nematode; worm;	
KW	Ostertagia ostertagi protein; immunogenic protein; gene; ss.	
XX		
OS	Ostertagia ostertagi.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	11..706
FT		/*tag= a
XX		/product= "Ostertagia ostertagi immunogenic protein"
PN	US2004052817-A1.	
XX		
PD	18-MAR-2004.	
XX		
PF	13-SEP-2002; 2002US-00243319.	
XX		
PR	13-SEP-2002; 2002US-00243319.	
XX		
PA	(GELD/) GELDHOF P.	
PA	(VERC/) VERCAUTEREN I.	
PA	(DMAE/) DE MAERE V.	
PA	(CLAER/) CLAERBOUT B.	
PA	(VERC/) VERCRUYSSSE J.	
XX		
PI	Geldhof P, Vercauteren I, De Maere V, Claerebout B, Vercruyssse J;	
XX		
DR	WPI; 2004-247704/23.	
DR	P-PSDB; ADK13847.	
XX		
PT	New Ostertagia ostertagi proteins and nucleic acid sequences, useful for	
PT	diagnosing, preventing or treating infections caused by Ostertagia	
PT	ostertagi in animals.	

XX Claim 5; SEQ ID NO 9; 36pp; English.  
XX  
CC The invention describes a nucleic acid sequence (I) encoding an  
CC Ostertagia osteragi protein, or part of the nucleic acid sequence that  
CC encodes an immunogenic fragment of the protein, where the nucleic acid  
CC sequence or its part has at least 85% homology with the nucleic acid  
CC sequence of the O. osteragi gene having any of the 7 fully defined  
CC sequences of 306-1761 base pairs, as given in the specification. The  
CC composition and methods are useful for diagnosing, preventing or treating  
CC infections caused by O. osteragi in animals. This sequence encodes a  
CC novel immunogenic Ostertagia osteragi protein.

XX  
SQ Sequence 763 BP; 223 A; 169 C; 191 G; 180 T; 0 U; 0 Other;

Query Match 100.0%; Score 763; DB 12; Length 763;  
Best Local Similarity 100.0%; Pred. No. 1.7e-231;  
Matches 763; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGAACTGCTATGTCGGCGCTGTTGTAGTGTCTCTCTGCGCCCTGTTCTCTATGCG 60  
DB 1 GAGAACTGCTATGTCGGCGCTGTTGTAGTGTCTCTCTGCGCCCTGTTCTCTATGCG 60  
OY 61 CGAAGCAGGCTTTGTTGTTCCGAATAGTCTAAGCCAAAGTGACAGCGCGAGCATTTT 120  
DB 61 CGAAGCAGGCTTTGTTGTTCCGAATAGTCTAAGCCAAAGTGACAGCGCGAGCATTTT 120  
OY 121 CCTCGATTTTCACAATGATGTTGTCGAAATATAGACATTGGAATGTTGATTAAGCTG 180  
DB 121 CCTCGATTTTCACAATGATGTTGTCGAAATATAGACATTGGAATGTTGATTAAGCTG 180  
OY 181 GACAGTAAATGACAGACGCGGTGATCTTGGTCCAGCTCAGAACATGTACAAAGTGACTG 240  
DB 181 GACAGTAAATGACAGACGCGGTGATCTTGGTCCAGCTCAGAACATGTACAAAGTGACTG 240  
OY 241 GGATTGCAACTTGAAGAAGTAGACACACACAGATTGCCCATGATCCCCCTAAC 300  
DB 241 GGATTGCAACTTGAAGAAGTAGACACACACAGATTGCCCATGATCCCCCTAAC 300  
OY 301 GATTAATACCAAGCTGGCTCAAAATATCGCTAGATGGCTGTAATTCAAGACAGTGAAGA 360  
DB 301 GATTAATACCAAGCTGGCTCAAAATATCGCTAGATGGCTGTAATTCAAGACAGTGAAGA 360  
OY 361 AGAGACAGTCTGCAACAAGTATCGGTATTTGGGTAGCGGCATCGCTGGATTATGA 420  
DB 361 AGAGACAGTCTGCAACAAGTATCGGTATTTGGGTAGCGGCATCGCTGGATTATGA 420  
OY 421 AGGACGAAACTTGACCAATTTGCTAACAGTGGGCTGAACCTCTAGCAAAATTGCAAA 480  
DB 421 AGGACGAAACTTGACCAATTTGCTAACAGTGGGCTGAACCTCTAGCAAAATTGCAAA 480  
OY 481 CTATAGAAACCGAAAGTTGATGTGCCATTAAGATCTGCCCCGCTCAGCAAAACATGTT 540  
DB 481 CTATAGAAACCGAAAGTTGATGTGCCATTAAGATCTGCCCCGCTCAGCAAAACATGTT 540  
OY 541 AGTATCCTGCGTGTATGGAAGCCCCCAACTTGCAACCGAAGTTATCTGGCAGGAAG 600  
DB 541 AGTATCCTGCGTGTATGGAAGCCCCCAACTTGCAACCGAAGTTATCTGGCAGGAAG 600  
OY 601 AAAAGCTTGTGTGTCGACGCTGCTCCAGATTCTTGTGTCGCAACCTGTGTGACAC 660  
DB 601 AAAAGCTTGTGTGTCGACGCTGCTCCAGATTCTTGTGTCGCAACCTGTGTGACAC 660  
OY 661 GCGAGATGTCGAGTGTTCGCCACCAAGTGTGCGCGTGGCCATGAAGGAAAGAAATT 720  
DB 661 GCGAGATGTCGAGTGTTCGCCACCAAGTGTGCGCGTGGCCATGAAGGAAAGAAATT 720  
OY 721 GGTAGTACCCCGAATAAATATTATCATGCAAAAAA 763  
DB 721 GGTAGTACCCCGAATAAATATTATCATGCAAAAAA 763

RESULT 2

ADK13844  
ID ADK13844 standard; cDNA; 828 BP.  
XX  
XX AC ADK13844;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
XX DE Ostertagia osteragi immunogenic protein cDNA seq id 7.  
XX  
XX KW antiparasitic; gene therapy; vaccine; nematode; worm;  
KW Ostertagia osteragi protein; immunogenic protein; gene; ss.  
XX  
XX OS Ostertagia osteragi.  
XX  
XX FH Key  
FH CDS Location/Qualifiers  
FT 11..721  
FT /tag= a  
FT /product= "Ostertagia osteragi immunogenic protein"  
FT /transl\_except= (pos:119..121, aa:Xaa)  
FT /note= "Xaa= Lys or Gln"  
FT /transl\_except= (pos:371..373, aa:Xaa)  
FT /note= "Xaa= Asn or Thr"  
XX  
XX PN US2004052817-A1.  
XX  
XX PD 18-MAR-2004.  
XX  
XX PF 13-SEP-2002; 2002US-00243319.  
XX  
XX PR 13-SEP-2002; 2002US-00243319.  
XX  
XX PA (GELD/) GELDHOF P.  
PA (VERC/) VERCAUTEREN I.  
PA (DMAE/) DE MAERE V.  
PA (CLAE/) CLAEREBOUT B.  
PA (VERC/) VERCRUYSSSE J.  
XX  
XX PI Geldhof P, Vercauteren I, De Maere V, Claerebout B, Vercruyssse J;  
XX  
XX DR WPI; 2004-247704/23.  
XX  
XX DR P-PSDB; ADK13845.  
XX  
XX PT New Ostertagia osteragi proteins and nucleic acid sequences, useful for  
PT diagnosing, preventing or treating infections caused by Ostertagia  
PT osteragi in animals.  
XX  
XX PS Claim 4; SEQ ID NO 7; 36pp; English.  
XX  
XX CC The invention describes a nucleic acid sequence (I) encoding an  
CC Ostertagia osteragi protein, or part of the nucleic acid sequence that  
CC encodes an immunogenic fragment of the protein, where the nucleic acid  
CC sequence or its part has at least 85% homology with the nucleic acid  
CC sequence of the O. osteragi gene having any of the 7 fully defined  
CC sequences of 306-1761 base pairs, as given in the specification. The  
CC composition and methods are useful for diagnosing, preventing or treating  
CC infections caused by O. osteragi in animals. This sequence encodes a  
CC novel immunogenic Ostertagia osteragi protein.  
XX  
XX SQ Sequence 828 BP; 255 A; 198 C; 179 G; 192 T; 0 U; 4 Other;

Query Match 35.8%; Score 273.4; DB 12; Length 828;  
Best Local Similarity 63.6%; Pred. No. 7.6e-76;  
Matches 447; Conservative 3; Mismatches 244; Indels 9; Gaps 2;

OY 38 CTCCTGCGCCCTGTTCTCTATGCGCGAAGCAGGCTTTGTTGTTGCCAATAGTCTAAGCCAA 97  
DB 44 CTGCTGTGCTGTGACATCAAAATACGAAAGCAGGTTTGTGCTGCCACAGATCTAAACCAA 103  
OY 98 AGTACAGCGCGAGCAGATTTCTCGATTTTTCACAAATGATGTTGTCGAATATAGCA 157  
DB 104 ACTGATGAGGCAAGMAAATCTTCTCGATTTTTCACAAATGATGTTGCGCGTGATATAGCA 163  
OY 158 CTTGAATGTTTGTATTAACCTGACAGTAAATGACAGCGCGTCA---TTCTTGTTCCA 214

```
Db 164 GGTGCAAGCCCGTGTCTCAACCTCACCAGCTGTTCAATGCGAATGTTCTCGGTCCA 223
Qy 215 GCTCAGACATGTACAAGTGAAGTGGATTGCACTTGAAGAGTAGACACACACAG 274
Db 224 GCTAAGACATGTACAGATGAGTGGACTGCAATCTGGAAGCAAAAGCAAGCAATG 283
Qy 275 ATTGGCCCATGCAATGATCCCTACCCATAAATACCAGCCTGGCTCAAAATATCGCTAGA 334
Db 284 ATTGGCCCATGCAATGACCTCTGCAATAGACAGATATTCACAAATCTCGCTCAR 343
Qy 335 TGGCTGTAATCAAAAGACAGTGAAGAGACAGTTCTGCAACAGTATCGTGTATTGG 394
Db 344 TGGCTGTAATTCACAAACAGTCAAGAAATGAAGTGTGACGCAACGCCCTGCTTGG 403
Qy 395 GTGAGCGCATCGCTGGATTATGAAGGACGAACTTGACCAATTGCTAACAGTGG 454
Db 404 GTACCGCATCACTACGAATCTTCACTGATACAGAGCTAACATTATTAATGCGCA 463
Qy 455 GCTGAACCTTAGCAAAACATTGCAAACTATAGAACCAGAAAGTTGATGCGCCATAAG 514
Db 464 ATTGACCACTATCCACATGCGAATGCGCAAAACCTAAAGTTGATGCTCACAAA 523
Qy 515 ATCTG-----CCCCGCTCAGCAAAACATGTTAGTATCTGCGTGTATGGAAGCCCAAA 568
Db 524 GTGTGCAAAATTCCTCCACCGGACAAATATGTTGTGCTTGTGCGTTATGCGCGCAAGTA 583
Qy 569 CTTCGACCGAAGCAAGTATCTGCGAAGAGAAAGGCTTGTGTGCGAAGCTGCTCCA 628
Db 584 CTTCAGATTAACGAAGTGTATGGAAGAGGACCAACTGCAATGCTATATCCC 643
Qy 629 GATTCAATCTGCTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTGTGCGCACAG 688
Db 644 AACTGCTTCTGCTGCAACAATCTGTGTGACACATAGCTGCTGCGACACTTGCAGACAG 703
Qy 689 TGTTCGGCGCTCGCCATGAAGCAAAAGAAATGTTAGTACCCC 731
Db 704 CCTTGAATCGACTTGAAGCAAAAGCGCTTGTGATGTCCC 746
```

RESULT 3

```
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
```

```
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
```

Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 5.4%; Score 41.4; DB 8; Length 2000;  
Best Local Similarity 9.3%; Pred. No. 0.066;  
Matches 54; Conservative 263; Mismatches 264; Indels 2; Gaps 1;

```
Qy 167 GGTTCATTAACCTGACAGTAAATGACAGCGGTCATTCTTGCTCAGCTCAGAACATG 226
Db 219 RKTGYTKMTTYSASRCMRAVYTTYSWACSSYTWCRSKRSMWMMKMRKMRKRYGWY 278
Qy 227 TACAAAGTGAAGTGGATGCACTTGAAGAGTAGACACAAACAGATTGCGCCATGC 286
Db 279 SWSYKMMCTAYKKSYSRWCTMYRGGWGATRYWGRGYSMAMMYKMYWRYGYKM 338
Qy 287 AATGATCCCTACCGATAAATACCAAGCTGCTCAAAATATCGCTAGATGCTGACTTC 346
Db 339 KRGWAGRMWMMRSMCRWSKACTYMRWMMRTRRRRMWAKSSRTSRKGRKMRKRYK 398
Qy 347 AAGACAGTGAAGAGACAGTCTGCAACAAGTATCGTGTATGTTGGTGAAGCGCATCG 406
Db 399 RMGYSRMRSCKRARWMTKRSRAWMMGRCGCMTCRMKSYGMWRKWSKMAKRYKMS 458
Qy 407 CTGGATTTATGAAGCGCACAACTTGACCAATTGCTAACAGTGGGCTGAACCTCTA 466
Db 459 RMTRWRKKCSRTTMMGKTGGMGTGRCRYKRSKGRKRRRRWGRMYRMWRYYM 518
Qy 467 GCAAACTTGCAAACTATAGAAACCGAAGTTGG--ATGTGCCATAGATCTGCCCG 524
Db 519 SARYTMRYCARKKYSYSAARXARCWYRGKYWAGMMWKYKRYMYKMMWYKRYKSK 578
Qy 525 CTCAGCAAAACATGTAGTATCTGCTGTATGGAAGCCCAACTTGACCGAAGCAAG 584
Db 579 SWYCKMSYASCKMSARKAGAKCKSKMSAWSKMSRSRKCRKCAKSSAKRYAMWG 638
Qy 585 TTAATCTGCAGGAAGAAAGCTGTGTGTGCGACGCTGCTCCAGATTCTGCTGCG 644
Db 639 MTSGRMSRMSKSYTCYMRKWSKMSCTCTMYMYMSKYIYAKGYSYWRYYRAWCMYMRWY 698
Qy 645 ACAACCTGTGTGACACGCGAGATGCTGCGAGTGTGCGCACAGTGTGCGCGTGCAT 704
Db 699 YRYRYSYMTYMAWYTSSTWAMTGMKYSGRYWTISWYKCKCSWKYRSMWYWSWMAKTW 758
Qy 705 GAAGCGAAAGAAATGTTAGTCAACCCGAAATAAATATTTCAT 747
Db 759 KMWRRYATRMWMMWYRYSMKYTWCTMGMWYMWYMWRTYMKGRY 801
```

RESULT 4

```
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
```



KW gene; ds.  
XX  
OS Oryza sativa.  
XX WO2003000898-A1.  
XX  
PD 03-JAN-2003.  
XX  
PF 22-JUN-2001; 2001WO-IB001105.  
XX  
PR 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
DR WPI; 2003-175290/17.  
XX  
PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
PS Claim 27; SEQ ID NO 5263; 899pp; English.  
XX  
XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.  
XX  
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;  
Query Match 4.9%; Score 37.4; DB 8; Length 2000;  
Best Local Similarity 9.1%; Pred. No. 1.2;  
Matches 38; Conservative 198; Mismatches 179; Indels 2; Gaps 1;  
QY 54 CCTATGCCGAGCAGGCTTTGTGTCGCAATAGTCTAAGCCAAAGTGACAGCGGAGGC 113  
DB 547 MCRYMGYMTTTSRSRMYYTGRYKARYTSKRBYMYKYRYCYWYYYYGMYMKCSYMMRYG 488  
QY 114 AGATTTTCCTCGATTTTTCACAATGATGTCGTAATATAGCACTTGGAATGTTGA 173  
DB 487 YCKACKKCCVAMCWAAYSGMMMYWYRKYSKWMRMSTKYMMSMWYKCRSMKYGAKGCTG 428  
QY 174 TAACTGACAGTAATAAGCAGACGCGGTCAATCTTGGTCAGCTCAGAACATGTACAAG 233  
DB 427 CKMWTYCSYGVNMYTYMGSYKYSRCYKVMYMYKGMWYMYMSAYSSMTWYYIYAKY 368  
QY 234 TGGACTGGGATTCGAATCTTGAAGAGTAGACACAACAGATTGCCCATGCAATGATC 293  
DB 367 WKYWKRRGTMSWTGSKYKKYCTWVCWYMKRCYRWKMRKTKTKYKRCYCWRYATCY 308  
QY 294 CCCTACCGATAAATACCAAGCCTGGCTCAAAATATCGTAGATGCGTG--TACTTCAAGA 351  
DB 307 WCCCYRKRGMWSRSMRTAGKWKMRSMWSRWSYWKMYKMMKSKYSYMSYGWARSSG 248  
QY 352 CAGTGAAGAGAGAGAGTCTTGCACAAGATATCGTATTTGGGTGAGCGCATCGCTGGG 411  
DB 247 TWSRSAAKRTYKGYSTSRRAKMRACRMYSACRRYSRTSYCGCSYCGSSRWKYMSKSCS 188  
QY 412 ATTATGAAGGACAGAACTTGACCAATTGCTAACAGTGGGCTGAACCTCTAGC 468  
DB 187 MRMTCCSSWCSCTYTCYGGAMCWCSCMSMYMGSCGCTYRGWKRSKYSMCKKYCSC 131

RESULT 5

ABD32730  
ID ABD32730 standard; DNA; 260803 BP.  
XX  
AC ABD32730;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human cancer-associated genomic DNA HD15-016.  
KW Human; ds; cancer-associated protein; gene; cytostatic; cancer;  
KW Leukemia; lymphoma; CAP.  
XX  
OS Homo sapiens.  
XX  
PN WO2004074320-A2.  
XX  
PD 02-SEP-2004.  
XX  
PF 17-FEB-2004; 2004WO-US004730.  
XX  
PR 14-FEB-2003; 2003US-00367094.  
PR 14-MAR-2003; 2003US-00388838.  
PR 15-APR-2003; 2003US-00417375.  
PR 13-JUN-2003; 2003US-00461862.  
PR 15-SEP-2003; 2003US-00663431.  
PR 15-DEC-2003; 2003US-00737318.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Morris DW, Malandro MS;  
XX  
DR WPI; 2004-652914/63.  
XX  
PT New isolated cancer-associated polynucleotides and polypeptides useful  
PT for diagnosing, preventing or treating cancers, especially lymphoma and  
PT leukemia, or in screening for agents that modulate cancer.  
XX  
PS claim 16; seqid 316; 310pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising at least 10  
CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
CC in the specification, or its complement. The nucleic acids encode cancer-  
CC associated proteins. Also included are an expression vector comprising  
CC the isolated nucleic acid cited above, a host cell comprising the above  
CC recombinant nucleic acid or expression vector, a microarray for detecting  
CC a cancer-associated (CA) nucleic acid comprising at least one probe  
CC comprising at least 10 contiguous nucleotides of any of the above-  
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
CC an open reading frame of a CA sequence selected from any of the 95  
CC polynucleotide sequences as mentioned in the specification, or its  
CC complement), an isolated antibody, (or its antigen binding fragment) that  
CC binds to the above polypeptide, a hybridoma that produces the above  
CC monoclonal antibody, a pharmaceutical composition comprising the above  
CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
CC cells (comprising the antibody cited above, methods for diagnosing cancer  
CC or for detecting the presence or absence of cancer cells in an  
CC individual, a method for inhibiting growth of cancer cells in an  
CC individual, a method for delivering a therapeutic agent to cancer cells  
CC in an individual, an electronic library comprising the above  
CC polynucleotide or polypeptide (or their fragments), methods of screening  
CC for anticancer activity or for a bioactive agent capable of modulating  
CC the activity of a CA protein (CAP), methods for detecting cancer  
CC associated with expression of a polypeptide in a test cell sample, a  
CC method for treating cancers and a method for inhibiting the expression of  
CC CA gene in a cell. The composition and methods are useful for detecting,  
CC diagnosing, preventing and treating cancers, especially lymphoma and  
CC leukemia. These may also be used in screening for agents that modulate  
CC cancer. The present sequence is a human CAP genomic sequence. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 260803 BP; 68597 A; 58832 C; 58504 G; 74728 T; 0 U; 142 Other;

```
Query Match      4.7%; Score 36; DB 13; Length 260803;
Best Local Similarity 45.7%; Pred. No. 38;
Matches 126; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

OY 103 CAGCGGAGGAGATTTCTCGATTTTCAATGATGTCGCGAAATATAGCACTGG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259473 CTGGGAAAGATACTATCCCGCACTTATCTGAAAATGATGCAATTTGGCAGCTAT 259532

OY 163 AAATGTTTGAATTAACGTGACAGATTAATGACAGCGGGTCATTCTTGTCAGCTCAGAA 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259533 GTATGAATGGGTATGTGCACTTATATGCTTTCACATGCTTATCATGTGTTAAGCCACTA 259592

OY 223 CATGTACAAGTGAAGTGGATGCACTTGAAGAAGTAGCAGACAACAGATTGCGCC 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259593 AAATTTGGGGGTCTTTGTTTACAGCAGCTAGCAAGCATTACCCTGACTGCTCAAC 259652

OY 283 ATGCAATGATCCCTTACCGATAAATACAGCCTGGCTCAAAATATCGCTAGATGCTGTA 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259653 ATGCACTGAGCATTTGACAAATCCAGCTATCAGTATTTCACCTTACCTTCTGATTGTA 259712

OY 343 CTTCAAGACAGTGAAGAGAGACAGTTCTGCAACA 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259713 CCACATTCATTAAAGAGGAGACAGAAATTCACACACA 259748

RESULT 6
ABS97663/c
ID ABS97663 standard; DNA; 1537 BP.
XX
AC ABS97663;
XX
DT 23-DEC-2002 (first entry)
XX
DE Human glutathione-S-transferase 12 (GST12) gene sequence.
XX
KW Human; de; cytochrome P450 A1; CYP4501A1; UGT2B4; MDR1;
KW cytochrome P450 A2; CYP4501A2; cytochrome P450 02E; CYP45002E1; LTF;
KW adrenergic receptor beta1; ADBR1; aryl hydrocarbon; AHR; MRP3; NR112;
KW aryl hydrocarbon receptor nuclear translocator; ARNT; cathepsin S; CTSS;
KW cyclooxgenase 2; COX2; diazepam binding inhibitor; DBI; haematological;
KW epoxide hydroxylase 2; EPHX2; 5-lipoxygenase activating protein; FLAP;
KW glutathione-S-transferase 12; GST12; histamine-N-methyl transferase;
KW HNMT; kallikrein 2; KLK2; nicotinamide-N-methyl transferase; NNMT;
KW NADPH quinone oxidoreductase 2; NQO2; sulfoltransferase thermolabile; STM;
KW UDP-glucuronosyl transferase 2B4; UDP-glucuronosyl transferase 2B7;
KW UGT2B7; UDP-glucuronosyl transferase; UGT2B15; urokinase receptor; uPA;
KW multidrug resistance 1; lactotransferrin; orphan nuclear receptor;
KW multidrug resistance associated protein 3; cancer; prostate;
KW acetylcholine muscarinic receptor; CHMR1; CHMR2; CHMR3; CHMR4; CHMR5;
KW altered drug metabolism; cardiovascular function; colorectal tumour;
KW central nervous system; pulmonary; immunological; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200257410-A2.
XX
PD 25-JUL-2002.
XX
PF 28-NOV-2001; 2001WO-US044838.
XX
PR 28-NOV-2000; 2000US-00724389.
XX
PA (DNAS-) DNA SCI LAB INC.
XX
PI Guida M, Hall J;
XX
DR WPI, 2002-698522/75.
XX
PT Isolated nucleic acid molecules having polymorphisms in known human genes
PT e.g. cytochrome p450 and cathepsin S useful as genetic linkage markers
PT for locating, identifying and characterizing the genes responsible for
```

```
PT disorder-related traits.
XX
PS Example 12; Page 321; 714pp; English.
XX
CC This invention relates to the sequence of an isolated nucleic acid
CC molecule comprising at least one base variation from that of a known
CC human cytochrome P450 A1 (CYP4501A1), cytochrome P450 A2 (CYP4501A2),
CC cytochrome P450 02E1 (CYP45002E1), adrenergic receptor beta1 (ADBR1),
CC aryl hydrocarbon (AHR), aryl hydrocarbon receptor nuclear translocator
CC (ARNT), cathepsin S (CTSS), cyclooxgenase 2 (COX2), diazepam binding
CC inhibitor (DBI), epoxide hydroxylase 2 (EPHX2), 5-lipoxygenase activating
CC protein (FLAP), glutathione-S-transferase 12 (GST12), histamine-N-methyl
CC transferase (NNMT), NADPH quinone oxidoreductase 2 (NQO2),
CC sulfoltransferase thermolabile (STM), UDP-glucuronosyl transferase 2B4
CC (UGT2B4), UDP-glucuronosyl transferase 2B7 (UGT2B7), UDP-glucuronosyl
CC transferase (UGT2B15), urokinase receptor (uPA), multidrug resistance 1
CC (MDR1), lactotransferrin (LTF), multidrug resistance associated protein 3
CC (MRP3), orphan nuclear receptor (NR112), or acetylcholine muscarinic
CC receptor 1, 2, 3, 4, or 5 (CHMR1, CHMR2, CHMR3, CHMR4 or CHMR5) sequence.
CC The polymorphisms in the human genes cited in the invention are useful as
CC genetic linkage markers for locating and characterizing the genes that
CC are responsible for specific traits within the genome and eventually
CC identifying the genes responsible for a variety of disorder-related
CC traits as a result of their e.g., overexpression, constitutive
CC expression, mutation or underexpression, which may be used in diagnosing
CC and/or treating the disorders. The nucleic acid molecules comprising the
CC polymorphic sequences contained in CYP4501A1, CYP4501A2, CYP4502E1,
CC ARNT, EPHX2, GST12, NNMT, NQO2, NR112, STM, UGT2B4, UGT2B7, UGT2B15, AHR,
CC MDR1 and/or MDR3 are useful for screening individuals for altered drug
CC metabolism. The polymorphic sequences contained in CYP4501A1, CYP4501A2,
CC AHR, MDR1 and/or MDR3 may also be used to screen individuals for
CC susceptibility to cancer. Polymorphic sequences in ADRB1 or CHMR2 are
CC used to screen for altered cardiovascular function, in COX2 for altered
CC susceptibility to colorectal tumours, in DBI or CHMR1 for altered central
CC nervous system function, in FLAP and HNMT for altered pulmonary,
CC immunological or haematological function, in KLK2 for altered serine
CC protease activity in the prostate, in LTF for altered immunological or
CC haematological function, in CHMR3, CHMR4 or CHMR5 for altered central and
CC peripheral nervous system function. The present sequence represents a
CC polymorphic DNA sequence of the invention.
XX
SQ Sequence 1537 BP; 394 A; 326 C; 303 G; 514 T; 0 U; 0 Other;

Query Match      4.6%; Score 35.4; DB 6; Length 1537;
Best Local Similarity 46.8%; Pred. No. 4.7;
Matches 111; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

OY 522 CCGCTCAGCAAAACATGTAATCTCTGCTGATGAAAGCCCAACTTGACCGAAG 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 913 CCAGCCAGCCCAACATGTTGAACCCCATCTCTACTTAAACCCCATCTCTACTTAAAAAA 854

OY 582 AAGTTATCTGCGAGGAAGAAAGCTTGTGTGCGAGCGCTGTCAGATTTCATTCTGCT 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 853 TAAATAATTAGCAGAGCGAGTGGCTGTGCTGCAATCCAGCTACTGTGAGACTGAG 794

OY 642 GCGACAACCTGTGTGACACGCGAGATGCTGCGAGTGTGCCAACAGTGTGCGGCTGCG 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 793 GCAGAGAATTGCTTAAATCCGGAGGCGGAGGTGCAATGAGCCAAGATTGCGGCACTG 734

OY 702 CATGAAGCGAAAAGAAATTTGTAATGATCAACCCGAAATTAATATTCATGCAAAAAAAA 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 733 CACTGTAGCCTGGGCGAGAGAGAGACTTATCTCAAAAAAAGAAAAA 677

RESULT 7
ABN96867/c
ID ABN96867 standard; DNA; 7769 BP.
XX
AC ABN96867;
XX
DT 13-AUG-2002 (first entry)
XX
```



DE Gene #3365 used to diagnose liver cancer.  
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
KW metastatic liver tumour; cytostatic; expression profile; disease state;  
KM disease progression; drug toxicity; drug efficacy; drug metabolism.  
XX  
OS Homo sapiens.  
XX  
PN WO200229103-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 02-OCT-2001; 2001WO-US030589.  
XX  
PR 02-OCT-2000; 2000US-0237054P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
XX  
DR WPI; 2002-426119/45.  
XX  
XX  
PT Diagnosing and detecting the progression of liver cancer, hepatocellular  
PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
PT level of expression of two or more genes in a liver tissue sample.  
XX  
PS Claim 1; SEQ ID NO 3365; 298bp; English.  
XX  
XX The invention relates to a novel method for diagnosing and detecting the  
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumour in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN93503-ABN97455 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 7769 BP; 2138 A; 1595 C; 1726 G; 2306 T; 0 U; 4 Other;  
  
Query Match 4.6%; Score 35.4; DB 6; Length 7769;  
Best Local Similarity 46.8%; Pred. No. 10;  
Matches 111; Conservative 0; Mismatches 126; Indels 0; Gaps 0;  
  
Qy 522 CCGCTCAGCAAAACATGTAGTATCTGCGTGTATGGAAGCCCAACTTGCACCGAAGC 581  
Db 7145 CCAGCCAGCCCAACATGTGAAACCCCACTCTACTAAAAACCCATCTCTACTAAAAAA 7086  
Qy 582 AAGTTATCTGCGAGGAAGAAAGCTGTGTGTGCGACGCTCGTCCAGATTCAATTCGTCT 641  
Db 7085 TAAAAAATTAGCCAGACGCGAGTGGCGGTGCTGCAATCCAGCTAATCTCGTAGACTGAG 7026  
Qy 642 GCGACAACCTGTGTGACACGCGAGATGCTGCGAGTGTTCGCCACAGTGTGCGCGTGC 701  
Db 7025 GCAGGAGAAATGCTTAATCCGGAGGCGGAGGTTGCAATGAGCCAAGATTGCCCACTG 6966  
Qy 702 CATGAAGCGAAAAAGAAATTGTTAGTCAACCCGAATAAATATTCATGCAAAAAAAA 758  
Db 6965 CACTGTAGCTGTGGCGCAGACAGAGAGACTTATCTCAAAAAAAGAAAAA 6909

WP ABD32627\_2 200001 310000  
WP ABD32627\_3 300001 410000  
WP ABD32627\_4 400001 421609  
ID ABD32627 standard; DNA; 421609 BP.  
XX  
AC ABD32627;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Mouse cancer-associated genomic DNA MD13-017.  
XX  
XX Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;  
KW Leukaemia; lymphoma; CAP.  
XX  
XX Mus musculus.  
OS  
XX  
PN WO2004074320-A2.  
XX  
PD 02-SEP-2004.  
XX  
PF 17-FEB-2004; 2004WO-US004730.  
XX  
PR 14-FEB-2003; 2003US-00367094.  
PR 14-MAR-2003; 2003US-00388838.  
PR 15-APR-2003; 2003US-00417375.  
PR 13-JUN-2003; 2003US-00461862.  
PR 15-SEP-2003; 2003US-00663431.  
PR 15-DEC-2003; 2003US-00737318.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Morris DW, Malandro MS;  
XX  
XX WPI; 2004-652914/63.  
XX  
DR  
XX  
XX New isolated cancer-associated polynucleotides and polypeptides useful  
PT for diagnosing, preventing or treating cancers, especially lymphoma and  
PT leukemia, or in screening for agents that modulate cancer.  
XX  
PS disclosure; seqid 148; 310pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising at least 10  
CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
CC in the specification, or its complement. The nucleic acids encode cancer-  
CC associated proteins. Also included are an expression vector comprising  
CC the isolated nucleic acid cited above, a host cell comprising the above  
CC recombinant nucleic acid or expression vector, a microarray for detecting  
CC a cancer-associated (CA) nucleic acid comprising at least one probe  
CC comprising at least 10 contiguous nucleotides of any of the above-  
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
CC an open reading frame of a CA sequence selected from any of the 95  
CC polynucleotide sequences as mentioned in the specification, or its  
CC complement), an isolated antibody, (or its antigen binding fragment) that  
CC binds to the above polypeptide, a hybridoma that produces the above  
CC monoclonal antibody, a pharmaceutical composition comprising the above  
CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
CC cells(comprising the antibody cited above, methods for diagnosing cancer  
CC or for detecting the presence or absence of cancer cells in an  
CC individual, a method for inhibiting growth of cancer cells in an  
CC individual, a method for delivering a therapeutic agent to cancer cells  
CC in an individual, an electronic library comprising the above  
CC polynucleotide or polypeptide (or their fragments), methods of screening  
CC for anticancer activity or for a bioactive agent capable of modulating  
CC the activity of a CA protein (CAP), methods for detecting cancer  
CC associated with expression of a polypeptide in a test cell sample, a  
CC method for treating cancers and a method for inhibiting the expression of  
CC CA gene in a cell. The composition and methods are useful for detecting,  
CC diagnosing, preventing and treating cancers, especially lymphoma and  
CC leukaemia. These may also be used in screening for agents that modulate  
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences



XX Sequence 421609 BP; 107858A; 90251C; 92730G; 119218T; 0U; 115520Other;  
SQ  
Query Match 4.5%; Score 34.6; DB 13; Length 110000;  
Best Local Similarity 50.3%; Pred. No. 69;  
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;  
QY 509 CATAGATCTGCCCCGCTCAGCAAAACATGTAGTATCTCTGCGTATGGAAGCCCAAA 568  
DB 65832 CCTAAGACATGGGGCCCGGTGGCGATAGGTTACCCCTATGACGGGTAGGCTGTGACATT 65773  
QY 569 CTTGCAACCGAAGCAAGTATCTGCGAAGGAAGGCTTGTGTGCGACGCTGCTCCA 628  
DB 65772 AGAGGAACGAACATAAAACAGAGCCACGCGGATGGCATATTGCAACAGCCTAGTCCA 65713  
QY 629 GATTCAATTCTGCTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTG 677  
DB 65712 GCCTCAATTATTATAAAACAACAGGGGAGATGTGGGAAGCAATGTG 65664

RESULT 9

ACN44398/c  
ID ACN44398 standard; DNA; 194883 BP.

AC ACN44398;  
XX  
DT 18-NOV-2004 (first entry)

DE Human genomic sequence hCG20056.

XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

OS Homo sapiens.

PN WO2003073826-A2.

PD 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006235.

PR 01-MAR-2002; 2002US-00087192.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW;

DR WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 826; Opp; English.

XX The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA coding  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published

XX Sequence 194883 BP; 41041 A; 50809 C; 53834 G; 49199 T; 0 U; 0 Other;

Query Match 4.5%; Score 34.6; DB 11; Length 194883;  
Best Local Similarity 50.9%; Pred. No. 92;  
Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 150 ATATAGCACTTGAATGTTGATTAAGTGAACAGTAAATGACAGCGGTCATTCTTG 209  
DB 79227 AGATGGACAGTGCCTGTAGTCCAGCTACTCAGAGAGAGAGGTGAGAGGATCTACTG 79168  
QY 210 GTCCAGCTCAGAACATGTACAAAGTGAAGTGGGATTGCACTTGGAAAGATGACGAC 269  
DB 79167 GAGCCAGAGAGTTCAGGCTGAGAGTGGCTATGATTGCACTGTGAAATAGCCACTGTAC 79108  
QY 270 AACAGATTGCGCCATGATGATCCCTACCGATTAATACC 310  
DB 79107 TCACGCTGGGCAACACAGCGAAACCTATCTAAGAAATTC 79067

RESULT 10

ABQ24108/c  
ID ABQ24108 standard; DNA; 1003 BP.

AC ABQ24108;

DT 12-JUL-2002 (first entry)

DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 10699.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
KW drug; side effect; cancer; central nervous system; cardiovascular;  
KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
KW SNP; cell differentiation; ds.

XX Homo sapiens.

PN WO200218632-A2.

PD 07-MAR-2002.

PF 01-SEP-2001; 2001WO-EP010074.

PR 01-SEP-2000; 2000DE-01043826.

PR 05-SEP-2000; 2000DE-01044543.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guetig D;

DR WPI; 2002-371829/40.

PT Determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
CC particularly by detecting mutations or single nucleotide polymorphisms  
CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
CC investigating cell differentiation. The method allows the methylation  
CC status of many C residues to be determined simultaneously. ABQ13410-  
CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
CC for determining the degree of cytosine methylation described in the  
CC disclosure of the invention

SQ Sequence 1003 BP, 89 A; 129 C; 313 G; 472 T; 0 U; 0 Other;
Query Match 4.5%; Score 34.4; DB 6; Length 1003;
Best Local Similarity 52.0%; Pred. No. 7.9;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 616 CGACGCTCGTCCAGATTCTGTGCTCGACAACTGTGTGACACGCGAGATGCTCGAG 675
Db 986 CGCGTCCCGTCGCAACCCCTTACCTACGCGACCGACTATTAACAAATACTTAAACTCCAA 927

QY 676 TGTTCGCCACCAAGTGTGCGCGTCGCCATGAAGCGAAAGAAATTGTAGTCACCCCGA 735
Db 926 CGCTCGCCCCGAACTTTAAACGACGAAAAAACCAATTAAAAACGAAACAAATCGAA 867

QY 736 TAAATATTTCATGCAAAAAA 763
Db 866 TACAATAAACCCGAAAAA 839

RESULT 11
ABQ24109
ID ABQ24109 standard; DNA; 1003 BP.
AC ABQ24109;
XX
XX 12-JUL-2002 (first entry)
DT
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 10700.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
XX WO200218632-A2.
PN
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
XX (EPiG-) EPIGENOMICS AG.
PA
XX
PI Olek A, Piepenbrock C, Berlin K, Guettig D;
XX
XX WPI; 2002-371829/40.
DR
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (1) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNPs); and (11) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation

CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 1003 BP, 472 A; 313 C; 129 G; 89 T; 0 U; 0 Other;
Query Match 4.5%; Score 34.4; DB 6; Length 1003;
Best Local Similarity 52.0%; Pred. No. 7.9;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 616 CGACGCTCGTCCAGATTCTGTGCTCGACAACTGTGTGACACGCGAGATGCTCGAG 675
Db 18 CGCGTCCCGTCGCAACCCCTTACCTACGCGACCGACTATTAACAAATACTTAAACTCCAA 77

QY 676 TGTTCGCCACCAAGTGTGCGCGTCGCCATGAAGCGAAAGAAATTGTAGTCACCCCGA 735
Db 78 CGCTCGCCCCGAACTTTAAACGACGAAAAAACCAATTAACAAACGAAACAAATCGAA 137

QY 736 TAAATATTTCATGCAAAAAA 763
Db 138 TACAATAAACCCGAAAAA 165

RESULT 12
ABD32806\_1
Continuation (2 of 7) of ABD32806 from base 100001 (Human cancer-associated genomic DNA ;
WP Sequence split into 7 fragments LOCUS ABD32806 Accession Abd32806
WP Fragment Name Begin End
WP ABD32806\_0 1 110000
WP ABD32806\_1 100001 210000
WP ABD32806\_2 200001 310000
WP ABD32806\_3 300001 410000
WP ABD32806\_4 400001 510000
WP ABD32806\_5 500001 610000
WP ABD32806\_6 600001 653458

Query Match 4.5%; Score 34.2; DB 13; Length 110000;
Best Local Similarity 58.3%; Pred. No. 93;
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 283 ATGCAATGATCCCTTACCGATAAATACCAAGCCTGGCTCAAAATATCGTAGATGGCTGTA 342
Db 43817 ATGGAATATTCAATTCTGATGAGTGACAGTGTCTGAATATATGGCCAGTTGGCAGGA 43876

QY 343 CTTCAAAAGACAGTGAAGAGAGACAGTCTGCAACAGTATCG 385
Db 43877 GGCCAATTCTTAATAGAAGCAATTCTGCTTCAAGAAATAG 43919

RESULT 13
ABL10900/c
ID ABL10900 standard; cDNA; 3457 BP.
XX
XX ABL10900;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 27182.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX
XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
DR  
XX P-PSDB; ABB66797.  
DR  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
PS  
XX Claim 1; SEQ ID NO 27182; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3457 BP; 972 A; 732 C; 768 G; 985 T; 0 U; 0 Other;

Query Match 4.5%; Score 34; DB 4; Length 3457;  
Best Local Similarity 51.3%; Pred. No. 19;  
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 122 CTCGATTTTCAATGATGTTGTCGAAATATAGCACTTGAATGTTTGATAACTGG 181  
DB 2931 CACGTATACCTCACTCAAGAGATCGTACTTAACCTTTTCGAGTAGTTGCATATGCAAG 2872  
QY 182 ACAGTAAATGCAGACGGCGTCAATCTTGTCAGCTCAGAACATGTACAAAGTGACTGG 241  
DB 2871 GACGTAAATCCGCCAGCGGCATTCCAATTCAGCGAAGTATTATGACCAATCGAAGA 2812  
QY 242 GATTGCACTTGGAAGAGTAGCAGACACAACA 275  
DB 2811 TATTGAAATCTGTAAACATTGCGACGCGACACACA 2778

RESULT 14

AB18190/c  
ID ABL18190 standard; DNA; 4092 BP.

XX ABL18190;  
AC  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 6043.  
XX  
KW Drosophila; developmental biology; cell signaling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.

XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS  
XX Claim 1; SEQ ID NO 6043; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 4092 BP; 1185 A; 876 C; 895 G; 1136 T; 0 U; 0 Other;

Query Match 4.5%; Score 34; DB 4; Length 4092;  
Best Local Similarity 51.3%; Pred. No. 21;  
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 122 CTCGATTTTCAATGATGTTGTCGAAATATAGCACTTGAATGTTTGATAACTGG 181  
DB 1046 CACGTATACCTCACTCAAGAGATCGTACTTAACCTTTTCGAGTAGTTGCATATGCAAG 987  
QY 182 ACAGTAAATGCAGACGGCGTCAATCTTGTCAGCTCAGAACATGTACAAAGTGACTGG 241  
DB 986 GACGTAAATCCGCCAGCGGCATTCCAATTCAGCGAAGTATTATGACCAATCGAAGA 927  
QY 242 GATTGCACTTGGAAGAGTAGCAGACACAACA 275  
DB 926 TATTGAAATCTGTAAACATTGCGACGCGACACACA 893

RESULT 15

ABL13492/c  
ID ABL13492 standard; cDNA; 7160 BP.

XX ABL13492;  
AC  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 34958.  
XX  
KW Drosophila; developmental biology; cell signaling; insecticide;  
KW pharmaceutical; gene; ss.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR P-PSDB; ABB69389.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Claim 1; SEQ ID NO 34958; 21pp + Sequence Listing; English.



XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (AB57337-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
SQ Sequence 7160 BP; 2075 A; 1468 C; 1473 G; 2144 T; 0 U; 0 Other;

Query Match 4.5%; Score 34; DB 4; Length 7160;

Best Local Similarity 51.3%; Pred. No. 28;

Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 122 CTCGATTTTCACATGATGTTCTCGAATATATAGCACTTGGAATGTTGATTAACCTGG 181  
Db 1046 CACGTATACCTCACTCAAGAGATCGTACTTAACCTTTTCGAGTAGTTGCATATGCAAG 987  
QY 182 ACAGTAAATGACAGCGCGTCAATCTGCTCCAGCTCAGAACATGTACAAAGTGACTGG 241  
Db 986 GACGTAAATCCGCCAGCGGCATTCCAATTCAAGCGAAAGTATTATGACCAATCGAAGA 927  
QY 242 GATTGCACTTGAAGAAGTAGCAGCAACAGCA 275  
Db 926 TATTGAAATCTGTAAACATTCGACGCGCAGCACA 893

Search completed: March 4, 2006, 07:12:51  
Job time : 594 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2006, 07:19:52 ; Search time 189 Seconds  
(without alignments)  
7176.082 Million cell updates/sec

Title: US-10-527-771-9

Perfect score: 763  
Sequence: 1 gagaactgctatgtcgcg...tcatgcaaaaaaaaaa 763

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
7: /cgn2\_6/ptodata/1/ina/PP\_COMB.seq: \*  
8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq: \*  
9: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.2	5.1	7218	2	US-08-232-463-14
2	34.6	4.5	601	3	US-09-949-016-45471
3	34.6	4.5	8297	3	US-09-949-016-14790
4	34.6	4.5	178884	3	US-09-949-016-12733
5	34.6	4.5	178884	3	US-09-949-016-13039
6	34	4.5	601	3	US-09-949-016-142991
7	34	4.5	1141	3	US-09-806-708B-22
8	34	4.5	265038	3	US-09-949-016-15779
9	33.4	4.4	106746	3	US-09-326-402C-1
10	33.4	4.4	106746	3	US-09-326-402C-12
11	32.6	4.3	5238	3	US-09-620-312D-351
12	32.2	4.2	1302	2	US-08-678-304-2
13	32	4.2	420	3	US-09-641-638-232
14	32	4.2	420	3	US-10-170-097-232
15	32	4.2	430	3	US-09-621-976-2446
16	32	4.2	609	3	US-09-621-976-287
17	31.6	4.1	808	3	US-08-651-136C-15
18	31.6	4.1	808	3	US-09-229-911A-15
19	31.6	4.1	808	3	US-10-007-521-15
20	31.6	4.1	1031	3	US-08-651-136C-19
21	31.6	4.1	1031	3	US-09-229-911A-19
22	31.6	4.1	1031	3	US-10-007-521-19
23	31.6	4.1	1048	3	US-08-651-136C-17
24	31.6	4.1	1048	3	US-09-229-911A-17

25	31.6	4.1	1048	3	US-10-007-521-17	Sequence 17, Appl
26	31.6	4.1	58273	3	US-09-949-016-14679	Sequence 14679, A
27	31.6	4.1	222452	3	US-09-949-016-12968	Sequence 12968, A
28	31.4	4.1	681	3	US-08-450-482B-124	Sequence 124, App
29	31.4	4.1	681	3	US-08-151-064D-122	Sequence 122, App
30	31.4	4.1	690	3	US-08-450-482B-126	Sequence 126, App
31	31.4	4.1	690	3	US-08-151-064D-124	Sequence 124, App
32	31.4	4.1	765	2	US-08-173-510B-102	Sequence 102, App
33	31.4	4.1	765	2	US-08-458-218-100	Sequence 100, App
34	31.4	4.1	765	2	US-08-450-497-102	Sequence 102, App
35	31.4	4.1	765	3	US-08-450-482B-102	Sequence 102, App
36	31.4	4.1	765	3	US-08-151-064D-100	Sequence 100, App
37	31.4	4.1	2879	3	US-10-104-047-1564	Sequence 1564, App
38	31.4	4.1	6804	3	US-09-949-016-16082	Sequence 16082, A
39	31.4	4.1	6804	3	US-09-949-016-16083	Sequence 16083, A
40	31.2	4.1	1003	3	US-09-453-702B-214	Sequence 214, App
41	31.2	4.1	1003	3	US-10-114-170-214	Sequence 214, App
42	31.2	4.1	1141	3	US-09-806-708B-22	Sequence 207, App
43	31	4.1	918	3	US-09-489-039A-207	Sequence 1, Appli
44	31	4.1	1320	2	US-08-419-414-1	Sequence 1, Appli
45	31	4.1	1324	2	US-08-330-272-1	Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHIEFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: BP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMTU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-Fls

US-08-232-463-14

Query Match 5.1%; Score 39.2; DB 2; Length 7218;

Best Local Similarity 7.7%; Pred. No. 0.12;

Matches 32; Conservative 197; Mismatches 185; Indels 0; Gaps 0;

```
OY 102 ACAGCCGAGCAGATTTCCTCGATTTCACATATGTTCTCGAAATATGCACTTG 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1505 AAAACGGCATGTAGCATCACTGTAATTACCTATCTATGCAAGTAGTTAAAGAGATAGAA 1446
OY 162 GAAATGTTGATTAACAGACAGTAATGACAGCGGCTCATTTCTTGCTCCAGCTCAGA 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1445 GAATTTGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1386
OY 222 ACATGTACAAAGTGAAGTGGATTCGAACTTGGAAGAGTAGACACACAGATTGCGC 281
    : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1385 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1326
OY 282 CATGCAATGATCCCTACCGATAAATACGAGCTGGCTCAAAATATCGTAGATGGCTGT 341
    : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1325 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1266
OY 342 ACTTCAAGAAGACAGTGAAGAGACAGTCTGCAACAAGTATCGTGATTTGGTGAGCG 401
    : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1265 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1206
OY 402 CATCGCTGGATTATGAAAGGACGAAACTTGACCAATTTGCTAACAGTGGGCTGAAC 461
    : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1205 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1146
OY 462 CTCTAGCAAAACATTGCAAACTATAGAAACGAAAGTTGATGTGCCCATAGA 515
    : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1145 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1092
```

## RESULT 2

US-09-949-016-45471

Sequence 45471, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 45471

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-45471

Query Match 4.5%; Score 34.6; DB 3; Length 601;

Best Local Similarity 50.9%; Pred. No. 1.2;

Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
OY 150 ATATAGCACTTGAAGTGTGTTGATTAACAGACAGTAATGACAGCGGCTCATTTCTTG 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 375 AGATGCAAGTGGCTGTAGTCCAGCTACTCAGAGAGAGAGAGATGATCACTG 434
OY 210 GTCCAGCTCAGAACATGTACAAAGTGAAGTGGATTCGAACTTGGAAGAGTAGACGAC 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 435 GAGCCCAAGAGTTCAAGGCTGAAGTCCGCTATGATTCATCTGTGAATAAGCACTGTAC 494
OY 270 AACAGATTGCGCCATGCAATGATCCCTACCGATAAATACC 310
```

```
DB 495 TCCAGCCTGGGCAACACAGCGAAACCTATCTAAGAAATC 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 3

US-09-949-016-14790/c

Sequence 14790, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14790

LENGTH: 8297

TYPE: DNA

ORGANISM: Human

US-09-949-016-14790

Query Match 4.5%; Score 34.6; DB 3; Length 8297;

Best Local Similarity 47.8%; Pred. No. 3.9;

Matches 100; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
OY 555 ATGAAGCCCCAAACTTGCAACCGAAGTTATCTGCGAGAGGAAAGGCTTGCTGT 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7059 ATGAAGAACCCTGTCTCTAATAAACAACAAATTAAGCTGGCGGTGTCATGTCT 7000
OY 615 GCGACGCTGTCAGATTCATTCTGCTGCGACAACCTGTGTGACAGCGAGATGTCGA 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6999 GTAATCCAGTACTCAGAGAGCTGAGCAGAGAAATCGCTTGAAGTGAAGAGCAGAG 6940
OY 675 GTGTTGCCACCAAGTGTGCGCGCTGCCATGAAGCGAAAGAAATGTAGTACCCCGA 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6939 TTGTGTTGAACCGAGATCAGCGCTTCCCTTCAGCTGGGCAACAGAGCAACTCCGT 6880
OY 735 ATAAATATTTCATGCACAAAAAATTTTATTTT 763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6879 CTCAAAAAATTTTATTTTATTTTATTTTATTTT 6851
```

## RESULT 4

US-09-949-016-12733/c

Sequence 12733, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12733

LENGTH: 178883

TYPE: DNA

ORGANISM: Human







```
FEATURE:
NAME/KEY: misc feature
LOCATION: 70647..70794
OTHER INFORMATION: homology with genset EST : A241850
FEATURE:
NAME/KEY: misc feature
LOCATION: 68648..68741
OTHER INFORMATION: homology with genset EST : A241850
FEATURE:
NAME/KEY: misc feature
LOCATION: 82208..82229
OTHER INFORMATION: homology with genset EST : A241850
FEATURE:
NAME/KEY: allele
LOCATION: 278
OTHER INFORMATION: 99-1601-278 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 402
OTHER INFORMATION: 99-1601-402 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 472
OTHER INFORMATION: 99-1601-472 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 2955
OTHER INFORMATION: 99-13801-100 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 12167
OTHER INFORMATION: 99-13806-166 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 12536
OTHER INFORMATION: 99-13799-376 : polymorphic base T or G
FEATURE:
NAME/KEY: allele
LOCATION: 17593
OTHER INFORMATION: 99-13798-297 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 17606
OTHER INFORMATION: 99-13798-284 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 22079
OTHER INFORMATION: 99-1602-200 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 28964
OTHER INFORMATION: 99-13794-186 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 29003
OTHER INFORMATION: 99-13794-147 : polymorphic base C or G
FEATURE:
NAME/KEY: allele
LOCATION: 31077
OTHER INFORMATION: 99-13812-384 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 31766
OTHER INFORMATION: 99-13805-313 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 34791
OTHER INFORMATION: 99-1587-281 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 45751
OTHER INFORMATION: 99-1582-430 : polymorphic base C or T
FEATURE:
```

```
NAME/KEY: allele
LOCATION: 49847
OTHER INFORMATION: 99-1585-465 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 49855
OTHER INFORMATION: 99-1585-457 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 49886
OTHER INFORMATION: 99-1585-426 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 49900
OTHER INFORMATION: 99-1585-412 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 49906
OTHER INFORMATION: 99-1585-406 : polymorphic base C or A
FEATURE:
NAME/KEY: allele
LOCATION: 49921
OTHER INFORMATION: 99-1585-391 : polymorphic base C or A
FEATURE:
NAME/KEY: allele
LOCATION: 49939
OTHER INFORMATION: 99-1585-373 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 50256
OTHER INFORMATION: 99-1585-55 : polymorphic base C or A
FEATURE:
NAME/KEY: allele
LOCATION: 54955
OTHER INFORMATION: 99-1607-373 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 64239
OTHER INFORMATION: 99-1577-105 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 65436
OTHER INFORMATION: 99-1591-235 : polymorphic base A or G
```

Query Match 4.4%; Score 33.4; DB 3; Length 106746;  
Best Local Similarity 51.7%; Pred. No. 30;  
Matches 76; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
QY 342 ACTCAAGACAGTGAAGAGACAGTCTGCAACAAGTATGCTGTTATGGTGAGCG 401
Db 91433 AGTCAAGCCAGTGAATTAACAAGCAAGTGCAAGAAATGACATGCTTCGGGTGAGCA 91374
QY 402 CATCGCTGGATTATGAAGGACGAAACTTGACCAATTGCTAACCAAGTGGCTGAAC 461
Db 91373 GACTACTGCAATGAAGAAGGACACATCTTGGCTTAGAAGAAACAGAACACAGAAT 91314
QY 462 CTCTAGCAACATTTGCAACTATAGAA 488
Db 91313 CCCTACAACCATCTGATGCTGCCGAA 91287
```

RESULT 10  
US-09-326-402C-12/c  
Sequence 12, Application US/09326402C  
Patent No. 6759192  
GENERAL INFORMATION:  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Bougueleret, Lydie  
APPLICANT: Chumakov, Ilya  
TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)  
FILE REFERENCE: GEN-T112XCI  
CURRENT APPLICATION NUMBER: US/09/326,402C  
CURRENT FILING DATE: 1999-06-04



```

; PRIOR APPLICATION NUMBER: 60/088,187
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/102,324
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 106746
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..68647
; OTHER INFORMATION: 5'regulation region
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 66647..68647
; OTHER INFORMATION: promoter
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 97156..106746
; OTHER INFORMATION: 3'regulation region
; FEATURE:
; NAME/KEY: exon
; LOCATION: 68648..68741
; OTHER INFORMATION: exon0
; FEATURE:
; NAME/KEY: exon
; LOCATION: 70647..70794
; OTHER INFORMATION: exon1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 82208..82296
; OTHER INFORMATION: exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 83613..83823
; OTHER INFORMATION: exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 85298..85417
; OTHER INFORMATION: exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 86389..86445
; OTHER INFORMATION: exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 87496..87522
; OTHER INFORMATION: exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 87650..87775
; OTHER INFORMATION: exon6bis
; FEATURE:
; NAME/KEY: exon
; LOCATION: 88295..88383
; OTHER INFORMATION: exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 89484..89649
; OTHER INFORMATION: exon8
; FEATURE:
; NAME/KEY: exon
; LOCATION: 92749..97155
; OTHER INFORMATION: exon9
; FEATURE:
; NAME/KEY: exon
; LOCATION: 92749..92883
; OTHER INFORMATION: exon9bis
; FEATURE:
; NAME/KEY: exon
; LOCATION: 95821..97155
; OTHER INFORMATION: exon9ter
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 70647..70794
; OTHER INFORMATION: homology with genset EST : A241850
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 68648..68741
; OTHER INFORMATION: homology with genset EST : A241850
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 82208..82229
; OTHER INFORMATION: homology with genset EST : A241850
; FEATURE:
; NAME/KEY: allele
; LOCATION: 278
; OTHER INFORMATION: 99-1601-278 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 402
; OTHER INFORMATION: 99-1601-402 : polymorphic base w= A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 472
; OTHER INFORMATION: 99-1601-472 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2955
; OTHER INFORMATION: 99-13801-100 : polymorphic base T or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 12167
; OTHER INFORMATION: 99-13806-166 : polymorphic base G or A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 12536
; OTHER INFORMATION: 99-13799-376 : polymorphic base T or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 17593
; OTHER INFORMATION: 99-13798-297 : polymorphic base T or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 17606
; OTHER INFORMATION: 99-13798-284 : polymorphic base T or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 22079
; OTHER INFORMATION: 99-1602-200 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 28964
; OTHER INFORMATION: 99-13794-186 : polymorphic base T or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 29003
; OTHER INFORMATION: 99-13794-147 : polymorphic base C or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 31077
; OTHER INFORMATION: 99-13812-384 : polymorphic base T or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 31766
; OTHER INFORMATION: 99-13805-313 : polymorphic base T or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 34791
; OTHER INFORMATION: 99-1587-281 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 45751
; OTHER INFORMATION: 99-1582-430 : polymorphic base C or T
```

```

1 FEATURE:
2 NAME/KEY: allele
3 LOCATION: 49847
4 OTHER INFORMATION: 99-1585-465 : polymorphic base T or C
5 FEATURE:
6 NAME/KEY: allele
7 LOCATION: 49855
8 OTHER INFORMATION: 99-1585-457 : polymorphic base T or C
9 FEATURE:
10 NAME/KEY: allele
11 LOCATION: 49886
12 OTHER INFORMATION: 99-1585-426 : polymorphic base G or A
13 FEATURE:
14 NAME/KEY: allele
15 LOCATION: 49900
16 OTHER INFORMATION: 99-1585-412 : polymorphic base G or A
17 FEATURE:
18 NAME/KEY: allele
19 LOCATION: 49906
20 OTHER INFORMATION: 99-1585-406 : polymorphic base C or A
21 FEATURE:
22 NAME/KEY: allele
23 LOCATION: 49921
24 OTHER INFORMATION: 99-1585-391 : polymorphic base C or A
25 FEATURE:
26 NAME/KEY: allele
27 LOCATION: 49939
28 OTHER INFORMATION: 99-1585-373 : polymorphic base G or A
29 FEATURE:
30 NAME/KEY: allele
31 LOCATION: 50256
32 OTHER INFORMATION: 99-1585-55 : polymorphic base C or A
33 FEATURE:
34 NAME/KEY: allele
35 LOCATION: 54955
36 OTHER INFORMATION: 99-1607-373 : polymorphic base T or C
37 FEATURE:
38 NAME/KEY: allele
39 LOCATION: 64239
40 OTHER INFORMATION: 99-1577-105 : polymorphic base A or G
41 FEATURE:
42 NAME/KEY: allele
43 LOCATION: 65436
44 OTHER INFORMATION: 99-1591-235 : polymorphic base A or G

```

```
Query Match      4.4%; Score 33.4; DB 3; Length 106746;
Best Local Similarity 51.7%; Pred. No. 30;
Matches 76; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
```

```
OY      342 ACTTCAAAAGACAGTGAGAAGAGACAGTTCTGCAACAAGTATCGTGTATTGGGTGAGCG 401  
| | | | | | | | | | | | | | | | | | | | |  
DB      91433 AGTCAAAGCCAGTGAATTACAAAGCAGTGCAAAAGAAATGACATGCTTCCGGGTGAGCA 91374  
  
OY      402 CATCGCTGGATTATTGAAAGGCACGAAACTGACCATTGCTTAACCAAGTGGGCTGAAC 461  
| | | | | | | | | | | | | | | | | | | | |  
DB      91373 GACTACTGCAAATGAAAAAAGAGCACATCCTTTGGCTTAGAAGAAAACAGAACACAGAAT 91314  
  
OY      462 CTCTAGCAAAACATTGCCAACTATAGAA 488  
| | | | | | | | | | | | | | | | | | | | |  
DB      91313 CCCTCACAACCATCTGATGTGCTGCCGAA 91287
```

```
RESULT 11  
US-09-620-312D-351  
Sequence 351, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyun  
APPLICANT: Chen, Rui-hong
```

```

; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunru
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 351
; LENGTH: 5238
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (278)..(1108)
US-09-620-312D-351

```

Query Match	4.3%;	Score 32.6;	DB 3;	Length 5238;
Best Local Similarity	57.3%;	Pred. No. 14;		
Matches	59;	Conservative	0;	Mismatches 44;
			Indels	0;
			Gaps	0;

```

QY      343  CTTCAAAGCAGTGAAGAAGACAGTCTTCGAACAAGTATCGTGTATTGGGTGAAGCC 4022
      |||||
Db      38   CATCAAGATGAAGAGAGACAGTCTTTCAGAGGTGTCAGTTTTCCTCCGACCC 97
      |||||

QY      403  ATCGCTGGATTATGAAGCGCAAGAACTTGACCAATTGCT 445
      |||||
Db      98  CCTGCCAGATTATTGACAAGGACACCAACCAATCAGCT 140
      |||||

```

RESULT 12  
US-08-678-304-2  
; Sequence 2, Application US/08678304  
; Patent No. 5663067  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Shuang-yong  
; APPLICANT: Xiao, Jian-ping  
; APPLICANT: Maunus, Robert B.  
; TITLE OF INVENTION: METHOD FOR CLONING AND  
; TITLE OF INVENTION: PRODUCING THE SAPI RESTRICTION  
; TITLE OF INVENTION: ENDONUCLEASE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: NEW ENGLAND BIOLABS, INC.  
; STREET: 32 TOZER ROAD  
; CITY: BEVERLY  
; STATE: MA  
; COUNTRY: US  
; ZIP: 01915  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/678,304  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
;

```
NAME: WILLIAMS, GREGORY D
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-927-5054
TELEFAX: 508-927-1705
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1299
US-08-678-304-2
```

```
Query Match 4.2%; Score 32.2; DB 2; Length 1302;
Best Local Similarity 54.7%; Pred. No. 9.7;
Matches 64; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
```

```
OY 163 AAATGTTTGTAACTGACAGTAATGACAGCGGTCATCTTGTCCAGCTCAGAA 222
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 645 AATGATATACATCAACCGAAGTTGATGTGCGCATTCATTTCTGATACGGTTCAAGC 704

OY 223 CATGTACAAGTGGACTGGATGCACTTGAAGAAGTAGCAGACACAGATTGC 279
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 705 AATGCTGACGATGTGAGCAGGTCCTTTGGAGCGGTAGCATCATTTATGTTTC 761
```

## RESULT 13

```
US-09-641-638-232
Sequence 232, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GENSET.051CPI
CURRENT APPLICATION NUMBER: US/09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 232
LENGTH: 420
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 198
OTHER INFORMATION: 10-94-198 : polymorphic base G or T
NAME/KEY: misc_binding
LOCATION: 178..197
OTHER INFORMATION: 10-94-198.mis1, potential
NAME/KEY: misc_binding
LOCATION: 199..218
OTHER INFORMATION: 10-94-198.mis2, potential complement
NAME/KEY: primer_bind
LOCATION: 1..19
OTHER INFORMATION: upstream amplification primer
```

```
NAME/KEY: primer_bind
LOCATION: 403..420
OTHER INFORMATION: downstream amplification primer, complement
NAME/KEY: misc_binding
LOCATION: 186..210
OTHER INFORMATION: 10-94-198 potential probe
NAME/KEY: misc_feature
LOCATION: 351
OTHER INFORMATION: n=a, g, c or t
US-09-641-638-232
```

```
Query Match 4.2%; Score 32; DB 3; Length 420;
Best Local Similarity 54.4%; Pred. No. 6.7;
Matches 62; Conservative 1; Mismatches 51; Indels 0; Gaps 0;
```

```
OY 87 GTCTAAGCCAAAGTGACAGCGGAGGAGATTTCCTCGATTTCACATGATGTCGTC 146
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 173 GTCCATTCCAAGTTGTAAATGCTAGKCTTTTCTTTTCTTTTCCATMAAAGACCATT 232

OY 147 GAAATATAGCACTTGAATGTTTGTATTAAGTGAACAGTAATGCAGACGCGG 200
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 233 AACTTAAGTGTTGTAAATGCTTGTAAAGCTGAGATCTAAATGGGACAGG 286
```

## RESULT 14

```
US-10-170-097-232
Sequence 232, Application US/10170097
Patent No. 6794143
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GEN-T114XC2D1
CURRENT APPLICATION NUMBER: US/10/170,097
PRIOR FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 232
LENGTH: 420
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 198
OTHER INFORMATION: 10-94-198 : polymorphic base G or T
NAME/KEY: misc_binding
LOCATION: 178..197
OTHER INFORMATION: 10-94-198.mis1, potential
NAME/KEY: misc_binding
LOCATION: 199..218
OTHER INFORMATION: 10-94-198.mis2, potential complement
NAME/KEY: primer_bind
LOCATION: 1..19
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer_bind
LOCATION: 403..420
OTHER INFORMATION: upstream amplification primer
```



OTHER INFORMATION: downstream amplification primer, complement  
FEATURE:  
NAME/KEY: misc\_binding  
LOCATION: 186..210  
OTHER INFORMATION: 10-94-198 potential probe  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 351  
OTHER INFORMATION: n=a, g, c o r t  
US-10-170-097-232

Query Match 4.2%; Score 32; DB 3; Length 420;  
Best Local Similarity 54.4%; Pred. No. 6.7;  
Matches 62; Conservative 1; Mismatches 51; Indels 0; Gaps 0;

QY 87 GTCTAAGCCAAAGTGACAGCGCGAGGAGATTTCCTCGATTTCACAATGATGTCGTC 146  
DB 173 GTCCATTCCAGTTGTAATGCTAGKCTTTTTCATTAATAAAGACCATTT 232  
QY 147 GAAATATAGCACTTGAATGTTGATTAACGTGACAGTAATGACAGCGG 200  
DB 233 AACTTAAGTGCTGTTAATGCTTTGTAAGCTGAGATCTAATGCGGACAAG 286

RESULT 15  
US-09-621-976-2446  
Sequence 2446, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621, 976  
NUMBER OF SEQ ID NOS: 2000-07-21  
SOFTWARE: Patent.pm  
SEQ ID NO 2446  
LENGTH: 430  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 63..254  
US-09-621-976-2446

Query Match 4.2%; Score 32; DB 3; Length 430;  
Best Local Similarity 60.2%; Pred. No. 6.7;  
Matches 53; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 GAGAACTGCTATGTCGGCGGCTGTTGATGTTCTCTCTGCGCCTGTTCTCTATGC 60  
DB 7 GTGAATCTCTGTGACTGCGAGTGGTGTACTGAAGGACCACTGGCCTTTTACACAATGG 66  
QY 61 CGAAGCAGGCTTTTGTGTCTCGAATAGT 88  
DB 67 CTAACAGCGATTCTTGTACGCAAAATT 94

Search completed: March 4, 2006, 07:29:57  
Job time: 190 secs

**This Page Blank (uspto)**





```

OY 241 GGATTGCAACTTGGAGAGTAGCAGCAGACAGATTGCCCATGCAATGATCCCTACC 300
    |||||
Db 241 GGATTGCAACTTGGAGAGTAGCAGCAGACAGATTGCCCATGCAATGATCCCTACC 300
OY 301 GATAAATACCAGCCTGGCTCAAAATATCGCTAGATGGCTGTACTTCAAGAAGAGTGAAGA 360
    |||||
Db 301 GATAAATACCAGCCTGGCTCAAAATATCGCTAGATGGCTGTACTTCAAGAAGAGTGAAGA 360
OY 361 AGAGACAGTCTGCAACAGATATCGTGTATTGGGTGAGCGCATGCTGGGATTATGAA 420
    |||||
Db 361 AGAGACAGTCTGCAACAGATATCGTGTATTGGGTGAGCGCATGCTGGGATTATGAA 420
OY 421 AGGCACGAAACTTGACCAATTTGCTAACCCAGTGGGTGAACCTCTAGCAAAACATTGCAA 480
    |||||
Db 421 AGGCACGAAACTTGACCAATTTGCTAACCCAGTGGGTGAACCTCTAGCAAAACATTGCAA 480
OY 481 CTATAGAAACCGAAAGGTTGATGTGCCCATTAAGATCTGCCCGCTCAGCAAAACATGGT 540
    |||||
Db 481 CTATAGAAACCGAAAGGTTGATGTGCCCATTAAGATCTGCCCGCTCAGCAAAACATGGT 540
OY 541 AGTATCCTGCGTGTATGGAAGCCCAAACTTGACCCGAAAGATTATCTGGCAGAGAG 600
    |||||
Db 541 AGTATCCTGCGTGTATGGAAGCCCAAACTTGACCCGAAAGATTATCTGGCAGAGAG 600
OY 601 AAAGCTGTGTGTGCGACGCTGTCAGATTCAATTCGTGCGACAACTGTGTGACAC 660
    |||||
Db 601 AAAGCTGTGTGTGCGACGCTGTCAGATTCAATTCGTGCGACAACTGTGTGACAC 660
OY 661 GCGAGATGCTGCGAGTGTTCGCCACCAAGTTGGCGGTGCGCATGAAGCAAAAGAAATT 720
    |||||
Db 661 GCGAGATGCTGCGAGTGTTCGCCACCAAGTTGGCGGTGCGCATGAAGCAAAAGAAATT 720
OY 721 GGTAATCAACCCGGAATTAATATTCATGCAAAAAA 763
    |||||
Db 721 GGTAATCAACCCGGAATTAATATTCATGCAAAAAA 763

```

RESULT 2

```

US-10-527-771-1
; Sequence 1, Application US/10527771
; Publication No. US20050271683A1
; GENERAL INFORMATION:
; APPLICANT: Universality Gent
; TITLE OF INVENTION: Osterlagia vaccine
; FILE REFERENCE: 2002-015
; CURRENT APPLICATION NUMBER: US/10/527,771
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US 10/243,319
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Osterlagia osterlagi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(721)
; US-10-527-771-1

```

```

Query Match 35.8%; Score 273.4; DB 8; Length 828;
Best Local Similarity 63.6%; Pred. No. 6,4e-71;
Matches 447; Conservative 3; Mismatches 244; Indels 9; Gaps 2;
OY 38 CTCCTGCCCCCTGTTCTCTATGCCCCGAGCAGGCTTTTGTGTCGAATAGTCTAAGCCAA 97
    |||||
Db 44 CTGGTGTGCTGATCAATAATACCGAAGCAGGTTTGTGCTGCCAGCAGATCTAACCAG 103
OY 98 AGTGACAGCGCAGGAGATTTTCTCTCGATTTCACATGATGTGTCGAATATATAGCA 157
    |||||
Db 104 ACTGATGAGGCAAGAAATCTTCTCGATTTCACATCAAGTTGCGCGTGATATAGCA 163
OY 158 CTTGGAATGTTTGAATACTGACAGTAATGACAGCGGCTCA--TTCTTGGTCCA 214

```

```

Db 164 GGTCAAGCCCCGTGTGCTCAACCTCACCGAGCTGTTCARATGCGAAATGTCTCGGTCCA 223
    |||||
OY 215 GCTCAGAACATGTACAAAGTGAAGTGGATGCACTTGAAAGAGTAGCAGCAACAG 274
    |||||
Db 224 GCTAAGAACATGTACAAAGTGAAGTGGATGCACTTGAAAGAGTAGCAGCAACAG 283
OY 275 ATTGCCCATGCAATGATCCCCCTACCGATAAATACCAAGCCTGGCTCAAAATATCGCTAGA 334
    |||||
Db 284 ATTGCCCATGCAATGATCCCCCTACCGATAAATACCAAGCCTGGCTCAAAATATCGCTCAR 343
OY 335 TGGCTGTACTTCAAAAGACAGTGAAGACAGTCTGCAACAGATATGCTGATTGG 394
    |||||
Db 344 TGGCTACTTTCCAAAACAGTCAAGAAATGAAAGTGTGACGCAACGCCCTGGTCTGG 403
OY 395 GTGAGCGCATCGCTGGGATTATGAAGGACGAACTTGACCAATTTGCTAACCAGTGG 454
    |||||
Db 404 GTAACCGCATCACTACGAATCTTCAACCTGATACAGAAAGCTAACATTATTAATGCGCAA 463
OY 455 GCTGAACCTCTAGCAAAACATTGCAAACTATGAACCGAAAGTTGATGTCGCCATAAG 514
    |||||
Db 464 ATTAGACCACTATCCAACTATGCGAAGTGGCAAAACCTAAAGTTGATGTCGCCATAAG 523
OY 515 ATCTG-----CCCCGCTCAGCAAAACATGTTAGTATCTCGCTGTATGAAAGCCCCAAA 568
    |||||
Db 524 GTGTCAAAATTCGCCACCGGACAAATATGTTGTGTCTTGGCTTATGCGCGGCAAGTA 583
OY 569 CTTGACCGAACGAAGTATCTGCGACAGAAAGAGGCTGTGTGTCGACGCTGTCCA 628
    |||||
Db 584 CTCGAAGATTAAGAAAGTGTATGGGACAAAGGACCACTTGATGTGCAATGCTTATCCC 643
OY 629 GATTCATTCTGCTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTTCGCCACAG 688
    |||||
Db 644 AACTGTTCTGCTGCAACAATCTGTGTGACAAATAGCTGCGACACTTGCAGAGCAG 703
OY 689 TGTTGCGCGTGGCCATGAAGCGAAAGAAATGTTGTTCAACC 731
    |||||
Db 704 CTTGTAATGACTTGAAGCGAAAGCGGCTGTGATGTCCC 746

```

RESULT 3

```

US-11-121-086-42
; Sequence 42, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 42
; LENGTH: 142303
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-42

```

```

Query Match 4.8%; Score 36.6; DB 12; Length 142303;
Best Local Similarity 56.1%; Pred. No. 13;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
OY 174 TAACTGACAGTAATGACAGCGGTCATTCTTGTGTCAGCTGAGAAATGTACAAAG 233
    |||||
Db 104467 TAATCTGAACAATTAAGCAATTAAGCTCTTATGCTGCTGTGCTCTATTAACG 104526
OY 234 TGGACTGGGATTCGAACCTTGAAGAGTAGCAGCAACAAGATTGGCCATGCAATGATC 293
    |||||
Db 104527 TGGACTGGGATTTGTAAGTGAAGAGGTTTATCTGCTCATTTCTTCTGCGCATGAAG 104586

```



```
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 625826
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-625826
```

```
Query Match      4.4%; Score 33.4; DB 6; Length 570;
Best Local Similarity 48.2%; Pred. No. 10;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```
QY      468 CAACATTGCAACTATAGAAACCGAAAGGTTGGATGTGCCCATAGATCTGCCCCGCTC 527
           |||||
Db      252 CTAGCACTTTGGAAGACCGGAGCGAGCATCACTTGAGGTCAAGAGCTCGAAGCCAGCC 193

QY      528 AGCAAAACATGTTAGTATCTGCGTGTATGGAGCCCCAACTTGCAACCGAAGTTA 587
           |||||
Db      192 TGGTCAACATGTGTAATCTGTCTCTACTAAAAGTACAAAGTTTAGCTGAAATAGCTT 133

QY      588 TCTGCAGAGGAAGAAAGGCTTGTGTGCGACGCTCGTCCAGATTCAATCTGCTGCGACA 647
           |||||
Db      132 GAACCCGGGAGCGGAGAGCTGCACTGAGATTGTGCTCACTGCATTCCAGCCGAGGC 73

QY      648 ACCTGTGTGACACGC 662
           |||||
Db      72 AACAGAGTGAGACAC 58
```

```
RESULT 8
US-09-925-065A-625827/c
; Sequence 625827, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 625827
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-625827
```

```
Query Match      4.4%; Score 33.4; DB 6; Length 570;
Best Local Similarity 48.2%; Pred. No. 10;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```
QY      468 CAACATTGCAACTATAGAAACCGAAAGGTTGGATGTGCCCATAGATCTGCCCCGCTC 527
```

```
Db      252 CTAGCACTTTGGAAGACCGGAGCGAGCATCACTTGAGGTCAAGAGCTCGAAGCCAGCC 193

QY      528 AGCAAAACATGTTAGTATCTGCGTGTATGGAGCCCCAACTTGCAACCGAAGTTA 587
           |||||
Db      192 TGGTCAACATGTGTAATCTGTCTCTACTAAAAGTACAAAGTTTAGCTGGAATAGCTT 133

QY      588 TCTGCAGAGGAAGAAAGGCTTGTGTGCGACGCTCGTCCAGATTCAATCTGCTGCGACA 647
           |||||
Db      132 GAACCCGGGAGCGGAGAGCTGCACTGAGATTGTGCTCACTGCATTCCAGCCGAGGC 73

QY      648 ACCTGTGTGACACGC 662
           |||||
Db      72 AACAGAGTGAGACAC 58
```

```
RESULT 9
US-10-995-561-13272/c
; Sequence 13272, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13272
; LENGTH: 149382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(149382)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
; US-10-995-561-13272
```

```
Query Match      4.4%; Score 33.4; DB 8; Length 149382;
Best Local Similarity 55.7%; Pred. No. 1.1e+02;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
```

```
QY      82 GAATAGTCTAAGCCAAAGTGACAGCGGAGGACAGATTTCCTCGATTTCACAATGATGT 141
           |||||
Db      2406 GAATGGCCTAAGCTAAATCATCTTAAATTAATTAAGTTTAAAGTAAAGCAGCA 23947

QY      142 TCGTCGAATATAGCACTTGAATGTTTGATAAATGACAGTAATGACAGAC 196
           |||||
Db      23946 TGAATCAATTATAGAACTAGAAAATGTTTCAGAGCAATAAATTATTACAGAC 23892
```

```
RESULT 10
US-09-925-065A-301865
; Sequence 301865, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
```



```

; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 301865
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-301865

```

```

Query Match 4.4%; Score 33.2; DB 6; Length 574;
Best Local Similarity 52.1%; Pred. No. 12;
Matches 74; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

```

QY 271 ACAGATTGGCCATGCAATGATCCCTACCGATAATACAGCCTGCTCAAAATATCGC 330
DB 221 ACTGAATCTGTAGGCAATTATACACATGATTAAGTATTGTGTAATAACATATCTA 280
QY 331 TAGATGGCTGTAAGACAGTGAAGAGACAGTTCGACCAAGTATCGTGA 390
DB 281 AACATACAATGATACAGTGAAGAAATGAATGATTAATAAATAAGTATACCTGTA 340
QY 391 TTGGGTGAGCGCATCGCTGGA 412
DB 341 TAGGGCAAGTACCATGATGGA 362

```

RESULT 11

```

; US-09-925-065A-625825/c
; Sequence 625825, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 625825
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-625825

```

```

Query Match 4.3%; Score 33; DB 6; Length 570;
Best Local Similarity 47.7%; Pred. No. 14;
Matches 93; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

```

```

QY 468 CAAACATTCGAACTATAGAAACCGAAGCTTGATGTCCTAAGATCTGCCCCGCTC 527
DB 252 CTAGCACTTTGGAAGACCGAGGAGGACGATCACTTGAAGTCAAGAGCTCGAGACGCG 193
QY 528 AGCAAAACATGTAATCTGCGGTATGAGAGCCCAACTGACCGAAGCAAGTTA 587
DB 192 TGGTCAACAYGTGAATCTGTCTACTAAAGTACAAAGTTAGCTGAATAGCTT 133
QY 588 TCTGCGAGGAAGAAAGCTTGTGTGCGAGCTGCTCAGATTCATCTGCTGCGACA 647
DB 132 GAACCCGGGAGCGGAGAGCTGCACTGAGATGTGCGCACTGCATTCCAGCCGAGC 73
QY 648 ACCTGTGTGACACGC 662

```

DB 72 AACAGGTGAGACAC 58

RESULT 12

```

; US-09-925-065A-193665/c
; Sequence 193665, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 193665
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-193665

```

```

Query Match 4.3%; Score 32.6; DB 6; Length 576;
Best Local Similarity 49.7%; Pred. No. 18;
Matches 83; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

```

```

QY 349 AGACAGTGAAGAGACAGTTCGCAACAAGTATCGTATTTGGGTGAGCGCATCGCT 408
DB 252 ACAGACTCACCCAGAAACATGCTGCACACAGTCCCTAGGTATTTGTTAATCCAGTCGA 193
QY 409 GGGATTATGAAGGACGAACTTGACCAATTTGCTAACAGTGGGCTGAACCTTAGC 468
DB 192 CTGATATCTAAATTCACCATGACACCATAGAGATGTAATCACACCTCTGAGTTTAC 133
QY 469 AACATTCGAACTATAGAAACCGAAGCTTGATGTCCTAAGATCTGCCCCGCTC 515
DB 132 AAAATCTACTATATGTTAGTTGACAGGTTGGGTATCCCTAGAGAGA 86

```

RESULT 13

```

; US-09-925-065A-745205
; Sequence 745205, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 745205

```







Db 301 GATAAATACCAGCCCTGGCTCAAAATATCGCTAGATGGCTGTACTTCAAGA CAGTGAAGA 360

QY 361 AGAGACAGTTCCTGCAACAAGTATCGTGTA TTGGGTAGCGCATCGCTGGGA TTTATGAA 420

Db 361 AGAGACAGTTCCTGCAACAAGTATCGTGTA TTGGGTAGCGCATCGCTGGGA TTTATGAA 420

QY 421 AGGCACGAACTTGACCAATTTGCTAA C CAGTGGGCTGAACCTCTAGCA A CATTGCAAA 480

Db 421 AGGCACGAACTTGACCAATTTGCTAA C CAGTGGGCTGAACCTCTAGCA A CATTGCAAA 480

QY 481 CTATAGAAACCGAAAGGTTGGATGTG C CCAATAAGATCTGCCCCGCTCAGCA A A A C ATG GT 540

Db 481 CTATAGAAACCGAAAGGTTGGATGTG C CCAATAAGATCTGCCCCGCTCAGCA A A A C ATG GT 540

QY 541 AGTATCCTGCGTGTATGGAAGCCCCA A A C TTGCA C CCGA A C G A A G T T A T C T G G C A G G A A G G 600

Db 541 AGTATCCTGCGTGTATGGAAGCCCCA A A C TTGCA C CCGA A C G A A G T T A T C T G G C A G G A A G G 600

QY 601 AAAGCCTTGTGTGTGCGCAGCTGCT C C A G A T T C A T T C T G C T G C G A C A A C C T G T G A C A C 660

Db 601 AAAGCCTTGTGTGTGCGCAGCTGCT C C A G A T T C A T T C T G C T G C G A C A A C C T G T G A C A C 660

QY 661 GCGAGATGCTGCGAGTGTTCGCCA C C A G T G T T G C G C G T G C C A T G A A G C G A A A G A A A T T 720

Db 661 GCGAGATGCTGCGAGTGTTCGCCA C C A G T G T T G C G C G T G C C A T G A A G C G A A A G A A A T T 720

QY 721 GGTAGTCA C C C C G A A T A A A T A T T C A T G C A A A A A A A A A A A A A A 763

Db 721 GGTAGTCA C C C C G A A T A A A T A T T C A T G C A A A A A A A A A A A A A A 763

## RESULT 2

```

US-10-243-319C-7
; Sequence 7, Application US/10243319C
; Publication No. US20040052817A1
GENERAL INFORMATION:
APPLICANT: Universiteit Gent
TITLE OF INVENTION: Ostertagia vaccine
FILE REFERENCE: Unigent Ostertagia
CURRENT APPLICATION NUMBER: US/10/243, 319C
CURRENT FILING DATE: 2002-09-13
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Ostertagia ostertagi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(721)
US-10-243-319C-7

```

Query Match	35.8%;	Score 273.4;	DB 7;	Length 828;
Best local Similarity	63.6%;	Pred. No. 1.9e-78;		
Matches 447;	Conservative	3;	Mismatches 244;	Indels 9;
				Gaps 2;

QY	38	CTCCCTGGCCCTGTTCTCTCTATGCGGAAGCAGCCTTTGTGTCCGAATAGCTAAGCCAA	97
Db	44	CTGGTCTGTGTGACATCAAAATACCGAAGCAGTTTGTGCTGCCAGCAGATTTAAACCA	103
QY	98	AGTGACAGCGCGAGGCAGATTTCTCGATTTCACAATGATGTTCTCGAAATATAGCA	157
Db	104	ACTGATGAGGCAAGAMAATCTTCTCGATTTTCACAATCAAGTTCGCCGTGATATAGCA	163
QY	158	CTTGGAAATGTTTGATATACTGACAGTAATGCAGACGCGTCA--TTCTTGTTCCA	214
Db	164	GGTGCAAGCCCCGTTGCTCAACCTCACCGGAGCTGTTCARATCGGAATGTTCTCGGTCCA	223
QY	215	GCTCAGAACATGTACAAAGTGGACTGGGATTGCACTTGGAGAAGAGTAGCAGCACACAG	274
Db	224	GCTAAGAACATGTACAGAAATGGACTGGGACTGCAATCTGGAGAAGCAAAAGCAAAAGGCATG	283

QY	275	ATTGCGCCATGC	AAATGATCCCTACCGATAAATAC	CGACCCCTGGCTCAAAATATCGCTAGA	334
					:
Db	284	ATTGGCCATG	CACCTACGCTCTGCCAATTAGACACAGATATTCACAAAATCTCGCTCAR		343
QY	335	TGGCTGTACTT	CAAGAAGACAGTGAAGAAGACAGTTCGCAACAAGTATCGTGATTGG	394	
Db	344	TGGCTACTTTT	CAAAAACAGTCAGAAAAMTGAAGTGTGACGCAACGCCCTGGTCTTGG	403	
QY	395	GTGAGCGCAT	CGCTGGATTATGAAAGCAGAACTTGACCAATTGGCTAACCACTGG	454	
Db	404	GTAACCGCAT	CACACTACGAATCTTCAACCTGATACAGAACTAACATTTAATACCTGGCAA	463	
QY	455	GCTGAACCT	CTAGCAAAACATTGCAAACTATAGAAACCGAAAGTTGATGTCCTCATAG	514	
Db	464	ATTAGACCA	CTATCCACACTTGGCAACTGGCAAAACCTAAAAAGTTGATGTGCTACAAA	523	
QY	515	ATCTG-----	CCCCGCTCAGCAAAACATGTAGTATCCTGCGTATGGAAGCCCCCAA	568	
Db	524	GTTGTGCAAA	TTCCCCAACCGGAGCAAAATATGTTGTCTGCGCTTATGGCGGGAAGTA	583	
QY	569	CTTGCAAC	CGAAGGATTATCTGGCAGGAAGAAAGGCTTGTGTGCGACGCTGCTCCA	628	
Db	584	CTCCAAAGAT	TAACGAAGTTGTATGGGACAAAGGACCAACTTGCAATGTCAATGTATCCC	643	
QY	629	GATTCAAT	TCTGCTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTGTGCGCAACAG	688	
Db	644	AACTCGTT	CTGCTGCAACAACTGTGTGACACAATAAGCTGCTGCGACACTTCGCAAGCAG	703	
QY	689	TGTTGCGCG	TGCGCCATGAAGCGAAAGAAATTGGTAGTCACCC	731	
Db	704	CCTTGTAAAT	CGACTTGAAAGCGAAAGGCGTGTGATGTATGCC	746	

### RESULT 3

```

US-10-388-838-84
; Sequence 84, Application US/10388838
; Publication No. US20040180344A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/388, 838
; CURRENT FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 260803
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(260803)
; OTHER INFORMATION: n = A,T,C or G
US-10-388-838-84

```

Query Match	4.7%	Score 36;	DB 8;	Length 260803;
Best Local Similarity	45.7%;	Pred. No. 34;		
Matches 126;	Conservative	0;	Mismatches 150;	Indels 0;
			Gaps	0;

Oy	103	CAGCGCAGGCGAGATTTTCTCGATTTTCACAATGATGTTGCTCGAAATATAGCACTTGG	162
Db	259473	CTGGGAAAAGATPACTATCCCACTTATCTGAAATGAATGCACAAATTTGGCAGCTAT	259532
Oy	163	AAATGTTTGATTAACCTGCAAGTAAATGCAGACGCGGTCATTCTTGGTCCAGCTCAGAA	222
Db	259533	GTAATGATGGGTATGTGCACTTATATGTCCTTACACATGCTTATCATGTGGTTAAGCCACTA	259592
Oy	223	CATGTACAAAGTGCACCTGGGATTTGCCAATTGGAAAGTAGCAGCACAACAGATTGGCC	282
Db	259593	AAATTGGGGGTGTTTGTTCACAGCAGCTAGCAAGCATTAACCTGACTGCATTGGCTAAC	259652
Oy	283	ATGCAATGATCCCTACCGATTAATACCAAGCTGGCTCAAAATATCGCTAGATGGCTGTA	342

Db 259653 ATGCACCTAGCATTTGACAAATCCAGCTATCAGTATTTCACCTTACCTTCTGATTGTA 259712  
Qy 343 CTTCAAGACAGTGAAGAGAGACAGTCTGCACA 378  
Db 259713 CCACATTCATTAGAGGACAGAAATTCACACA 259748

## RESULT 4

US-09-925-065A-501602  
; Sequence 501602, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 501602  
; LENGTH: 591  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-501602

Query Match 4.6%; Score 34.8; DB 4; Length 591;  
Best Local Similarity 56.2%; Pred. No. 3.5;  
Matches 63; Conservative 1; Mismatches 48; Indels 0; Gaps 0;

Qy 322 AAATATCGCTAGATGGCTGACTTCAAGACAGTGAAGAGACAGTCTGCACAAGT 381  
Db 215 AAATATAGAGATGCTGTATGCTAATGAGAACTAGAAACTACTGCAGAGGA 274  
Qy 382 ATCGGTATTTGGTGAGCGCATCGCTGGATTATGAAGGACGAAACTT 433  
Db 275 AAAGGATTAATGCTGTGTAATGACCTGAGTAYAGAGGGCCCTTCCTT 326

## RESULT 5

US-10-741-601-5625/c  
; Sequence 5625, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: Cargill, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CU001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5625  
; LENGTH: 16977  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-5625

Query Match 4.5%; Score 34.6; DB 7; Length 16977;  
Best Local Similarity 47.8%; Pred. No. 24;  
Matches 100; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 555 ATGAGGCCCAACTTGCACCCGACGAAGTATCTGGCAGAGAGAAAGCCTGTGTCT 614  
Db 11730 ATGAGAACCTGTCTCTACTTAAACACAAATTAGCTGGCGTGTGTGTCT 11671  
Qy 615 GCGACGCTGCTCAGATTCTGCTGCGACCACTGTGTGACACGCGAGATGTCGGA 674  
Db 11670 GTAATCCAGTACTCAGAGGCTGAGGACAGAGATCGCTTGAATGAAGAGCAGAG 11611  
Qy 675 GTGTGCGCACCAAGTGTGCGCGTGCATGAAGCGAAAGAAATGTAGTCACCCCGA 734  
Db 11610 TTGTGTGAACCGAGATCAGCGCTTCCCTTCAGCCTGGGCAACAGAGCGAACTCCGT 11551  
Qy 735 ATAAATATTCATGCACAAAAAATTTT 763  
Db 11550 CTCACAAAAAATTTT 11522

## RESULT 6

US-10-087-192-826/c  
; Sequence 826, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 826  
; LENGTH: 194883  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-087-192-826

Query Match 4.5%; Score 34.6; DB 5; Length 194883;  
Best Local Similarity 50.9%; Pred. No. 85;  
Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 150 ATATGCACTTGAATGTTGATTAACCTGACAGTAATGACAGCGGTCTTCTT 209  
Db 79227 AGATGGACGCTGCTGTAGTCCAGCTACTCAGAGAGAGAGGTGAGAGATCACTG 79168  
Qy 210 GTCAAGCTCAGACATGTACAAAGTGAAGTGGATTGCACTTGAAGAAGTAGCAGAC 269  
Db 79167 GAGCCAGAGAGTCAAGGCTGAAGTGGCTATGATTCATCTGTGAATAAGCAGCTGTAC 79108  
Qy 270 AACAGTTGCGCCATGCAATGATCCCTTACCGATTAATACC 310  
Db 79107 TCCAGCTGGGCAACACAGCGAAACCTTATCTAAGAAATC 79067

## RESULT 7

US-10-367-094-122/c  
; Sequence 122, Application US/10367094  
; Publication No. US20040170982A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer  
; FILE REFERENCE: 529452001500  
; CURRENT APPLICATION NUMBER: US/10/367,094  
; CURRENT FILING DATE: 2003-02-14  
; NUMBER OF SEQ ID NOS: 203  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122  
; LENGTH: 421609

```
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(421609)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-122
```

```
Query Match          4.5%; Score 34.6; DB 7; Length 421609;
Best Local Similarity 50.3%; Pred. No. 1.3e+02;
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
```

```
QY 509 CATAAGATCTGCCCCGCTCAGCAAAACATGTAATCTCGTGTATGAAAGCCCCAAA 568
    |||||
DB 65832 CCTAAGACATGGGGCCGGTGGCGATAGGTTACCTTATGACGGGTAAGGCTGTGACATT 65773
    |||||
QY 569 CTGCAACCGAAGCAAGTATCTGGCAGAGAAAGGCTTGTGTGCGACGCTGCTCCA 628
    |||||
DB 65772 AGAGGAACGAACATAAAACAGAGCCACGGCGGATGGGCAATAATGCACAGGCTAGTCCA 65713
    |||||
QY 629 GATTCAATCTGCTGCGACACCTGTGTGACACGCGAGATGCTGCGAGTG 677
    |||||
DB 65712 GCCTCATTTTATTAAACAAACAGGGGAGATGTGGGAAGCCACATGTG 65664
    |||||
```

## RESULT 8

```
US-10-363-345A-10699/c
; Sequence 10699, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; NUMBER OF SEQ ID NOS: 2003-03-03
; SEQ ID NO 10699
; LENGTH: 1003
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 10699
US-10-363-345A-10699
```

```
Query Match          4.5%; Score 34.4; DB 8; Length 1003;
Best Local Similarity 52.0%; Pred. No. 6.2;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
```

```
QY 616 CGAGCTCGTCCAGATTCATTCTGCTGCGACAACCTGTGTGACACGCGAGATGCTGCGAG 675
    |||||
DB 986 CGCGTCCCGTCCGCAACCCCTACCTACGCGACGACTATACAAATACTAAACTCCAAA 927
    |||||
QY 676 TGTTCGCCACCAAGTGTGCGCGTCCGCAATGAAGCAAAAGAAATTGGTAGTCACCCCGAA 735
    |||||
DB 926 CGCTCGCCCCGAACTTTAAACGACGAAAAAAACCAATAAAAAACGAAAAATAATCGAA 867
    |||||
QY 736 TAAATATTTCATGCACAAAAA 763
    |||||
DB 866 TACAATAAACCCGAAAAA 839
    |||||
```

## RESULT 9

```
US-10-363-345A-10700
; Sequence 10700, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
```

```
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 10700
; LENGTH: 1003
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 10700
US-10-363-345A-10700
```

```
Query Match          4.5%; Score 34.4; DB 8; Length 1003;
Best Local Similarity 52.0%; Pred. No. 6.2;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
```

```
QY 616 CGAGCTCGTCCAGATTCATTCTGCTGCGACAACCTGTGTGACACGCGAGATGCTGCGAG 675
    |||||
DB 18 CGCGTCCCGTCCGCAACCCCTACCTACGCGACGACTATACAAATACTAAACTCCAAA 77
    |||||
QY 676 TGTTCGCCACCAAGTGTGCGCGTCCGCAATGAAGCAAAAGAAATTGGTAGTCACCCCGAA 735
    |||||
DB 78 CGCTCGCCCCGAACTTTAAACGACGAAAAAAACCAATAAAAAACGAAAAATAATCGAA 137
    |||||
QY 736 TAAATATTTCATGCACAAAAA 763
    |||||
DB 138 TACAATAAACCCGAAAAA 165
    |||||
```

## RESULT 10

```
US-10-363-483A-10699/c
; Sequence 10699, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 10699
; LENGTH: 1003
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 10699
US-10-363-483A-10699
```

```
Query Match          4.5%; Score 34.4; DB 9; Length 1003;
Best Local Similarity 52.0%; Pred. No. 6.2;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
```

```
QY 616 CGAGCTCGTCCAGATTCATTCTGCTGCGACAACCTGTGTGACACGCGAGATGCTGCGAG 675
    |||||
DB 986 CGCGTCCCGTCCGCAACCCCTACCTACGCGACGACTATACAAATACTAAACTCCAAA 927
    |||||
QY 676 TGTTCGCCACCAAGTGTGCGCGTCCGCAATGAAGCAAAAGAAATTGGTAGTCACCCCGAA 735
    |||||
DB 926 CGCTCGCCCCGAACTTTAAACGACGAAAAAAACCAATAAAAAACGAAAAATAATCGAA 867
    |||||
QY 736 TAAATATTTCATGCACAAAAA 763
    |||||
DB 866 TACAATAAACCCGAAAAA 839
    |||||
```

## RESULT 11





APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 51217  
LENGTH: 388  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_53630C.1  
US-10-437-963-51217

Query Match 4.5%; Score 34.2; DB 7; Length 388;  
Best Local Similarity 51.7%; Pred. No. 4.4;  
Matches 78; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 611 GTGTGCGACGCTCGTCCAGATTCATTCTGCTGCGACAACCTGTGTGACACCGGAGATGCT 670  
DB 155 GCGTTCGAGGGTGGTGGCGACGCCCCCGCGCGATAGCGGAGGGCGCACGACGCGCC 96  
QY 671 GCGAGTGTTCGCCACCAAGTGTTCGCGCTGCCCATGAAGCGAAGAAATTGGTAGTCACC 730  
DB 95 GCGAGTGTTCGCCACCGCCACCGCACTCCAGTCCACCGCAAGCGGATAGACCA 36  
QY 731 CCGAATAAATATTCATGCAGAAAAAAA 761  
DB 35 AAGAAAAATATATAAATATCAATAAGAAAA 5

RESULT 15

US-10-461-862-4  
Sequence 4, Application US/10461862  
Publication No. US20050090434A1  
GENERAL INFORMATION:  
APPLICANT: David W. Morris  
APPLICANT: Marc S. Malandro  
TITLE OF INVENTION: Novel Therapeutic Targets in Cancer  
FILE REFERENCE: 529452001800  
CURRENT APPLICATION NUMBER: US/10/461,862  
CURRENT FILING DATE: 2003-06-13  
NUMBER OF SEQ ID NOS: 184  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 653458  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(653458)  
OTHER INFORMATION: n = A,T,C or G  
US-10-461-862-4

Query Match 4.5%; Score 34.2; DB 9; Length 653458;  
Best Local Similarity 58.3%; Pred. No. 2.2e+02;  
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 283 ATGCAATGATCCCTACCGATAATACCACTGGCTCAAAATATCGTAGATGCTGTA 342  
DB 143817 ATGGAATATTCACTTCTGATGAGTGACAGTGTCTGAATATATGCGCAGTGCGACAGA 143876  
QY 343 CTTCAAGACAGTGAAGAGACAGTCTGCAACAAGTATCG 385  
DB 143877 GGCCAATTCTTAATAGAAGACCAATTCTGCTTCAAGAATAG 143919

Search completed: March 4, 2006, 07:49:31  
Job time : 827 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2006, 07:13:05 ; Search time 3916 Seconds  
(without alignments)  
9116.074 Million cell updates/sec

Title: US-10-527-771-9

Perfect score: 763

Sequence: 1 gagaaactgtatgtcgcgcg.....tcatacaaaaaaaaaaaaaa 763

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	478.4	62.7	766	6	CB036994 TC_ad2_28
2	476.8	62.5	757	6	CB038073 TC_ad2_41
3	475.8	62.4	757	6	CB036437 TC_ad2_21
4	474.2	62.1	760	6	CB038292 TC_ad2_44
5	474.2	62.1	766	6	CB038112 TC_ad2_42
6	473.6	62.1	756	6	CB036391 TC_ad2_21
7	473.2	62.0	755	6	CB036975 TC_ad2_28
8	472.6	61.9	751	6	CB038323 TC_ad2_44
9	472.6	61.9	754	6	CB038204 TC_ad2_43
10	470.2	61.6	771	6	CB038345 TC_ad2_44
11	469.2	61.5	733	6	CB037619 TC_ad2_36
12	469.2	61.5	763	6	CB037329 TC_ad2_32
13	467.6	61.3	739	6	CB036489 TC_ad2_32
14	466.2	61.1	800	6	CB037895 TC_ad2_39
15	466.2	61.1	758	6	CB036762 TC_ad2_25
16	464.6	60.9	721	3	BM052055 TC_ad2_02
17	464.6	60.9	752	6	CB036281 TC_ad2_19
18	464.6	60.9	757	6	CB036396 TC_ad2_21
19	464.6	60.9	783	6	CB036875 TC_ad2_27
20	464.2	60.8	736	6	CB037631 TC_ad2_36
21	463	60.7	768	6	CB037578 TC_ad2_35
22	463	60.7	778	6	CB038400 TC_ad2_45

23	462.2	60.6	777	6	CB038358 TC_ad2_44
24	462	60.6	783	6	CB038266 TC_ad2_43
25	461.4	60.5	757	6	CB038393 TC_ad2_45
26	459.2	60.2	704	6	CB037577 TC_ad2_35
27	457.8	60.0	728	6	CB036543 TC_ad2_23
28	457.4	59.9	711	6	CB038828 TC_ad2_50
29	457	59.9	759	6	CB037859 TC_ad2_39
30	456.2	59.8	709	6	CB036361 TC_ad2_20
31	454.6	59.6	739	6	CB036890 TC_ad2_27
32	454.4	59.6	704	6	CB038110 TC_ad2_42
33	454.4	59.6	720	6	CB036228 TC_ad2_19
34	454	59.5	767	6	CB036381 TC_ad2_21
35	453.4	59.4	704	6	CB038648 TC_ad2_48
36	453.2	59.4	756	6	CB038256 TC_ad2_43
37	453	59.4	743	6	CB038839 TC_ad2_50
38	452.8	59.3	735	6	CB036573 TC_ad2_23
39	451.6	59.2	734	6	CB037846 TC_ad2_38
40	451.2	59.1	728	6	CB037691 TC_ad2_37
41	450.8	59.1	695	6	CB039176 TC_ad2_54
42	450.2	59.0	703	6	CB036788 TC_ad2_26
43	450.2	59.0	760	6	CB037386 TC_ad2_33
44	449.4	58.9	700	6	CB038612 TC_ad2_48
45	449.4	58.9	743	6	CB037372 TC_ad2_33

#### ALIGNMENTS

RESULT 1  
CB036994  
LOCUS  
DEFINITION  
CB036994 766 bp mRNA linear EST 15-JAN-2003  
Tc\_ad2\_28F03\_TEXF1 Teladorsagia circumcincta adults library 2  
Teladorsagia circumcincta cDNA clone Tc\_ad2\_28F03 5' similar to  
CAD56659 Ancylostoma-secreted protein-like prote Osteragia, mRNA  
sequence.  
ACCESSION  
CB036994 GI:27756239  
VERSION  
CB036994.1 GI:27756239  
KEYWORDS  
EST.  
SOURCE  
Teladorsagia circumcincta  
ORGANISM  
Teladorsagia circumcincta  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhaditida; Strongylida;  
Trichostrongyloidea; Haemonchidae; Osteragiinae; Teladorsagia.  
REFERENCE  
1 (bases 1 to 766)  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guiliano, D.,  
Hall, N., Quayle, M. and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
TITLE  
JOURNAL  
COMMENT  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3UT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCRFL  
BACKWARD: T7PL  
Plate: 28 row: F column: 03  
Seq primer: TEXF1  
High quality sequence stop: 519.  
location/Qualifiers  
1. 766  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="Tc\_ad2\_28F03"  
/sex="mixed"



/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site\_1: BamHI; Site\_2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match	62.7%;	Score 478.4;	DB 6;	Length 766;	
Best Local Similarity	79.8%;	Pred. No. 1.1e-133;			
Matches 579;	Conservative 0;	Mismatches 141;	Indels 6;	Gaps 1;	
QY	19	GGCTGTGTAGTGTGCTGTTCTCTGCCCCCTGTTCTCTATGCCCCGAAGCAGGCTTTGTTG	78		
DB	25	GGCAGTCACAGTTGCTGTTCTCTGCTTGTCTCTCCATGCCGAAGCAGGCTTTGTTG	84		
QY	79	TCCGAATAGTCTAAGCCAAAGTGACAGCGGAGGAGATTTCTCGATTTTCACAAATGA	138		
DB	85	CCCAAAACTCTAAGCCAACTGACACGCGAGGAGATTTCTCGATTTTCACAAATGA	144		
QY	139	TGTTGCTCGAATATAGACACTTGGAAATGTTTGATAAATCGACAGTAATGACAGCGC	198		
DB	145	TATTCGTGGAAACATAGCACTTGGAAAGAGCTTGTAAGT-----TCAGACTGATCC	198		
QY	199	GGTCATTTCTTGCTCCAGCTCAGAACATGTACAAAGTGAAGTGGGATTGCACTTGAAGA	258		
DB	199	GGTTGTTCTTGCTCCAGCTCAGAACATGTACAAAGTGAAGTGGGATTGCACTTGAAGA	258		
QY	259	AGTAGCAGCACACAGATTGCCCATGCAATGATCCCCCTACCGATTAATACCAAGCTGGC	318		
DB	259	AAAAGCAGCTCAACAGATTGGCAATGCAAGGTTCTCTACCAATGATCCCAAGCTGGC	318		
QY	319	TCAAAATATCGCTAGATGGCTGTAATTCAAAAGCAGTGAAGAAGACAGTTCTGACA	378		
DB	319	TCAAAATATCGCAAGATGGCTGTAATGCAACAGTGAAGAAGACAGTTTGAACA	378		
QY	379	AGTATCGTGTATTGGTGAGCGCATCGCTGGATTATGAAGGACGAAACTTGACCA	438		
DB	379	AGTACCGTGTCTTGAGTAAACCCCATCGCTACGATTCATGAAGGACGCACTTGATCG	438		
QY	439	ATTGCTAACCAAGTGGGCTGAACCTCTAGCAAAACATTTGCAAACTATAGAAACGAAAGGT	498		
DB	439	TTTTGCTAACCAAGTGGGCTGAACCGCTTGCCCAACATTTGCTAACTGAAAAAATCGGAAGGT	498		
QY	499	TGGATGTGCCCATTAAGATCTGCCCCGCTCAGCAAAACATGTTAGTATCTCTGCTGATGG	558		
DB	499	TGGATGTGCCCTAACAGATCTGCCCGCTGAAAAACATGTTAGTATCTCTGCTGATGG	558		
QY	559	AAGCCCCCAACTTGCAACCGAAGTTATCTGGCAGGAAGAAAGGCTTGTGTGCGA	618		
DB	559	AAGCCAAAACTTTCACCAAGCAAGTAAATTTGGAGCAAGAGTACTTGCGAGTGCAA	618		
QY	619	CGCTGTCAGATTCATTCTGCTGCGCAACCTGTGTGACACGCGAGATGCTGCGAGTGT	678		
DB	619	CTCTTATCCGATTCGTTTGTGTGTGACAGCTGTGTGACACGCAATGAGCTGCGAGCT	678		
QY	679	TGCGCACCAAGTGTGCGCGCTGCGCATGAAGCGAAAGAAATTTGTTAGTCAACCCGAATAA	738		
DB	679	TGCGCACCAAGTGTGTGTGATGACATGAGGCGAAGAAATTTCTGTTAATCTTGAGCAA	738		
QY	739	AATATT 744			
DB	739	AATATT 744			

RESULT 2  
CB038073 757 bp mRNA linear EST 15-JAN-2003  
LOCUS  
DEFINITION  
Teladorsagia circumcincta adults library 2  
Teladorsagia circumcincta cDNA clone Tc\_ad2\_41E11 5' similar to  
CAD56659 Ancylostoma-secreted protein-like prote Ostertagia, mRNA  
sequence.  
ACCESSION  
CB038073

VERSION  
CB038073.1 GI:27757318  
KEYWORDS  
EST.  
SOURCE  
Teladorsagia circumcincta  
ORGANISM  
Teladorsagia circumcincta  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongylidae; Haemonchidae; Oostertaglinae; Teladorsagia.  
REFERENCE  
1 (bases 1 to 757)  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Gulliano, D.,  
Hall, N., Quayle, M., and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
COMMENT  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCTRF1  
BACKWARD: T7PL  
Plate: 41 row: B column: 11  
Seq primer: TEXP1  
High quality sequence start: 5  
High quality sequence stop: 519.  
Location/Qualifiers

FEATURES

source  
1. 757  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="Tc\_ad2\_41E11"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site\_1: BamHI; Site\_2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match	62.5%;	Score 476.8;	DB 6;	Length 757;	
Best Local Similarity	79.6%;	Pred. No. 3.4e-133;			
Matches 578;	Conservative 0;	Mismatches 142;	Indels 6;	Gaps 1;	
QY	19	GGCTGTGTAGTGTGCTGTTCTCTGCCCCCTGTTCTCTATGCCGAAGCAGGCTTTGTTG	78		
DB	25	GGCAGTCACAGTTGCTGTTCTCTGCTTGTCTCTCCATGCCGAAGCAGGCTTTGTTG	84		
QY	79	TCCGAATAGTCTAAGCCAAAGTGACAGCGGAGGAGATTTCTCGATTTTCACAATGA	138		
DB	85	CCCAAAACTCTAAGCCAACTGACACGCGAGGAGATTTCTCGATTTTCACAATGA	144		
QY	139	TGTTGCTCGAATATAGCACTTGGAAATGTTTGATATACTGGAACAGTAATGACAGCGC	198		
DB	145	TATTCGTGGAACATAGCACTTGGAAAGAGCTTGTAAGT-----TCAGACTGATCC	198		
QY	199	GGTCATTTCTTGCTCCAGCTGGAACATGTACAAAGTGAAGTGGGATTGCACTTGAAGA	258		
DB	199	GGTTGTTCTTGCTCCAGCTGGAACATGTACAAAGTGAAGTGGGATTGCACTTGAAGA	258		
QY	259	AGTAGCAGCACACAGATTGCCCATGCAATGATCCCCCTACCGATTAATACCAAGCTGGC	318		
DB	259	AAAAGCAGCTCAACAGATTGGCGAATGCAAGGTTCTCTACCAATAGATCCCAAGCTGGC	318		
QY	319	TCAAAATATCGCTAGATGGCTGTAATTCAAAAGCAGTGAAGAAGACAGTTCTGACA	378		
DB	319	TCAAAATATCGCAAGATGGCTGTAATGCCCAACAGTGAAGAAGACAGTTTGAACA	378		

QY	379	AGTATCGTGGTATTTGGGTGAGCGCATCGCTGGGATTTATGAAGGCA	CGAAACTTGACCA	438	
Db	379	AGTACCCTGGTCTTGGGTAA	CCCATCGCTACGATTTATGAAGGCA	CGGCACCTTGATCG	438
QY	439	ATTGCTAACCACTGAGTGGGCTGAACCTTAGCAAA	CATTGCAAACTATAGAAACCGAAAGT	498	
Db	439	TTTGTAAACCACTGGGCTGAACCGCTTGCCAAC	ATTGCTAACTGAAAAAATCGAAAGT	498	
QY	499	TGATGTGCCCATTAAGATCTGCCCGCTCAGCAAAA	CATGTAGTATCCTCGGTATGG	558	
Db	499	TGATGTGCTTACAAAATCTGCCCGCCTGAAAAA	CATGTGTATCTCTGTGTATGG	558	
QY	559	AAGCCCAAACTTGCAACCGAAGAAATATCTGCGAGAAAGAAAGCTTGTGTGCGA	618		
Db	559	AAGCAAAAACCTTCAACCAACGAAGTAATTTGGAGCAAGAAAGTACTTGGAGTGCAA	618		
QY	619	CGCTCGTCCAGATTCAATCTGTGCGACAACCTGTGTGACACCGCAGATGCTGCGAGTGT	678		
Db	619	CTCTTATCCGGATTCTGTTTGTGTGTGACACGCTGTGTGACACCGCATGGAGCTGCGAGCCT	678		
QY	679	TCGCCAACCACTGTGCGCGCTGCCCATGAAGCGAAAAAGAAATTGTTAGTCAACCCCGAATTA	738		
Db	679	TCGCCAACCACTGTGTGTGATGCACTGAGCGGAGAAAAATTCTGTTAACATCTTGAGCAA	738		
QY	739	AATATT	744		
Db	739	AATATT	744		

RESULT 3  
CB036437  
LOCUS  
DEFINITION  
CB036437 757 bp mRNA linear EST 15-JAN-2003  
Tc ad2\_21G05\_TEXF1 Teladorsagia circumcincta adults library 2  
Teladorsagia circumcincta cDNA clone Tc ad2\_21G05 5' similar to  
CAD56659 Ancylostoma-secreted protein-like prote Ostertagia, mRNA  
sequence.  
ACCESSION  
VERSION CB036437 GI:27755682  
KEYWORDS  
SOURCE EST.  
ORGANISM Teladorsagia circumcincta  
Teladorsagia circumcincta  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongyloidea; Haemonchidae; Ostertagiinae; Teladorsagia.  
REFERENCE  
AUTHORS 1 (bases 1 to 757)  
Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D.,  
Hall,N., Quayle,M. and Barrell,B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
COMMENT Contact: Blaxter ML

```

FEATURES
source
Location/Qualifiers
1. .757
/organism="Teladorsagia circumcincta"
/mol_type="mRNA"
/db_xref="taxon:45464"

```

```

/clone="Tc_ad2_21G05"
/sex="mixed"
/dev_stage="adult"
/clone_lib="Teladorsagia circumcincta adults library 2"
/notes="Vector: pTriplEx2; Site_1: BamHI; Site_2: BamHI;
Teladorsagia circumcincta is a parasitic nematode. The
library was constructed from mRNA from Teladorsagia
circumcincta mixed adults."

```

Query Match	62.4%	Score 475.8	DB 6	Length 757
Best Local Similarity	79.1%	Pred. No. 6.8e-133		
Matches 580	Conservative 0	Mismatches 147	Indels 6	Gaps 1
QY 19	GGCTGTGTAGTGTCTGTTCTCTGCGCCCTGTTCTCTATGCGGAGCAGGCTTTGTG	78		
DB 25	GGCAGTCACAGTGGTGTCTCTGCGCTTTGTCTCTATGCGGAGCAGGCTTTGTG	84		
QY 79	TCCGATAGTCTAAGCCAAAGTGCAGCGCGAGGAGATTTCTCGATTTTCAATGA	138		
DB 85	CCCAAAACTCTAAGCCAACTGACACAGCGAGGAGATTTCTCGATTTTCAATGA	144		
QY 139	TGTTGTCGAAATATAGCACTTGGAATGTTGATTAACCTGACAGTAAATGCAGCG	198		
DB 145	TATTGTCGGAACATAGCACTTGGAAGAGCTGTGTAAC-----TCAAGACTGATCC	198		
QY 199	GGTCATTCTTGTCAGCTCAGAACATGTACAAAGTGACTGGGATTGCAACTTGAAGA	258		
DB 199	GGTGTCTTGGTCCAGCTCAGAACATGTACAAAGTGACTGGGACTGTGACCTGAACA	258		
QY 259	AGTAGCAGCACAACAGATTGCGCCATGCAATGATCCCTACCATAAATACAGCCTGGC	318		
DB 259	GAAAGCAGCTCAACAGATTGCGCAATGCACAGGTTCTTACCAATGATCCAGCCTGGC	318		
QY 319	TCAAATATTCGTAGATGGCTGTACTTCAAGACAGTGAAGAGAAGACAGTTCTGACACA	378		
DB 319	TCAAATATTCGCAAGATGGCTGTATATGCAACAGTGAAGAGAAGACAGTTTGAAGAG	378		
QY 379	AGTATCGTGTATTGGGTGAGCGCATCGCTGGATTATGAAGGACGAACTTGACCA	438		
DB 379	AGTACCGTGTCTTGGGTAAACCCCATCGCTACGATTCAATGAAGGACAGGCACTTGATCG	438		
QY 439	ATTGCTAACCAAGTGGCTGAACCTCTAGCAAAACATTGCAAACTATAGAAACCGAAAGGT	498		
DB 439	TTTCGCTAACCAAGTGGCTGAACCGCTTGCAACATTTGTAAGTGAAGAAATCGAAGGT	498		
QY 499	TGGATGTGCCCATAGATCTGCCCCGCTCAGCAAAACATGTAATCTTGCCTGCTATGG	558		
DB 499	TGGATGTGCTTACAGATCTGCCCCGCTGAAAAACATGTCGTATCTGTGTATGG	558		
QY 559	AAGCCCCAAACTTGCACCGGAAAGAAAGTTATCTGCGAGGAAGAAAGCCTTGTGTGCGA	618		
DB 559	AAGCCAAAACTTCAACCAACGAAGTAATTGGGAGCAAGGAAGTAATTGCGAGTGCAA	618		
QY 619	CGCTGTCAGATTCATTCTGCTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTG	678		
DB 619	CTCTTATCCGGAATGTTTGTGCTGTGACACAGCCTGTGTGACACGCGTGAAGCTGCGAGCCT	678		
QY 679	TGCGCACCAAGTGTGCGCGTGCAGCATGAAGCGAAAGAAATGGTAGTCAACCCGAATAA	738		
DB 679	TGCGCACCAAGTGTGCTGATGCATGAGGCGGAGGAAGAAATCTGTTAAACATCTTGAGCAG	738		
QY 739	AATATTGATGCAA 751			
DB 739	AATATTGATTTAA 751			

RESULT 4	LOCUS	DEFINITION
CB038292	760 bp mRNA	EST 15-JAN-2003
CB038292	Tc ad2_44A12	TEXF1 Teladorsagia circumcincta adults library 2
	Teladorsagia circumcincta cDNA clone Tc ad2_44A12 5'	similar to
CAD56659	Ancyclostoma-secreted protein-like prote	Ostertagia, mRNA

	sequence.
ACCESSION	CB038292
VERSION	CB038292.1 GI:27757537
KEYWORDS	EST.
SOURCE	Teladorsagia circumcincta
ORGANISM	Teladorsagia circumcincta Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida; Trichostrongyloidea; Haemonchidae; Overtaglinae; Teladorsagia. 1 (bases 1 to 760)
REFERENCE	Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D., Hall,N., Quayle,M. and Barrell,B. Edinburgh University/Sanger Centre Nematode EST Project Unpublished (2000) Contact: Blaxter ML Institute of Cell, Animal and Population Biology University of Edinburgh Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK. Tel: +44 131 650 6760 Fax: +44 131 670 5450 Email: mark.blaxter@ed.ac.uk The library was prepared by Dr Diane Redmond and Dr David Knox, Moredun Research Institute, Midlothian, UK. PCR was performed by Ye Jieru, ICAPB, University of Edinburgh. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).
AUTHORS	PCR Primers FORWARD: TEXPCCRF1 BACKWARD: T7PL Plate: 44 row: A column: 12 Seq primer: TEXF1 High quality sequence stop: 517. Location/Qualifiers
JOURNAL	1..760
COMMENT	/organism="Teladorsagia circumcincta" /mol_type="mRNA" /db_xref="taxon:45464" /clone="Tc ad2_44A12" /sex="mixed" /dev_stage="adult" /clone_lib="Teladorsagia circumcincta adults library 2" /note="Vector: pTriplEx2; Site_1: BamHI; Site_2: BamHI; Teladorsagia circumcincta is a parasitic nematode. The library was constructed from mRNA from Teladorsagia circumcincta mixed adults."
FEATURES	
source	
ORIGIN	
Query Match	62.1%; Score 474.2; DB 6; Length 760;
Best Local Similarity	79.0%; Pred. No. 2.le-132;
Matches 579; Conservative	0; Mismatches 148; Indels 6; Gaps 1;
OY	19 GGCTGTGTAAGTTGCTGTTCTTCCTGGCCTGTTCTCTATGCCGAAGCAGGCTTTGTTG 78
Db	24 GGCAGTCATAAGTTGCTGTTCTTGCTTTGCTCCTCCATGCGAGCAGGCTTTGTTG 83
OY	79 TC CGAATAGTCTAA GCCCAAAGTAGCAGCGCGAGGCAGATTTCCTCGATTTCACAATGA 138
Db	84 CCCAAAACCTTAAGCCAAC TGACCA GCAGGCAGATTTTCCTCATTTTCACAATGA 143
OY	139 TGTTCGTCGAAATATAGCACCTTGGAAATGTTGATTAACCTGACAGTAATG CAGACG 198
Db	144 CATTCGTCGGAACATAGCACTTGGAAGA GCTTGTA CT-----TCACGACTGATCC 197
OY	199 GGTCAATCTTGCTCCAGCTCAGAACATGTACAAAGTGGA CTGGGATTGC AACCTGGA GA 258
Db	198 GGTGTTCTTGCTCCAGCTCAGAACATGTACAAAGCTGGA CTGGGACTGTGACC TTGAACA 257
OY	259 AGTAGCAGCACAACAGATTGCGCCATGCATGATCCCCTACCGATAAATAC CAGCCTGGC 318
Db	258 AAAAGCAGCTCAACAGATTGCGCAATGCAGCGTT CCTTACCAATAGATCC CAGCCTGGC 317
OY	319 TCAAAATATCGCTAGATGGCTGTACTTCAAAGACAGTGAAGAGAGACAGT TCTGCAACA 378

Db	318	TCAAATATCGCAAGATGGCTGTATTATGCGCAACAGTGAAGAAGACAAGGTTTGGAGCA	377
QY	379	AGTATCGGTATTGGGTGAGCGCATCGCTGGATTATGAAAGGCACGAACTTGAACA	438
Db	378	AGTACCGTGTCTTGGGTAGCCCCATCGCTACGATTCATGAAGGGCACGGCACTTGATCG	437
QY	439	ATTTGCTAACCAGTGGGCTGAACCTTAGCAAACATTCGAAACTATAGAAACCGAAAGGT	498
Db	438	TTTTGCTAACCAGTGGGCTGAACCGCTTGCCAACATGTCTAACTGAAAAATCGGAAGGT	497
QY	499	TGGATGTGCCCATAGACTTGCCTCCCGCTCAGCAAAACATGCTAGTATCCTGCGTATGG	558
Db	498	TGGATGTGCTTACAGACTTGCCTCCCGCTGAAAAACATGGTCGTATCCTGTGTATGG	557
QY	559	AAGCCCCAAACTTGCACCGAACGAAGTTATCTGGCAGAGAAGAAAGCCTTGTGTGCGA	618
Db	558	AAGCCAAAACTTTCACCAAAAGAAATTAATTGGGAGCAAGGAATACTTGGAGTGCAA	617
QY	619	CGCTCGTCCAGATTCATCTGCTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTGT	678
Db	618	CTCTTATCCGATTCGTTTGTGCTGTGACAGCCTGTGTGACACGAATGGAAGCTGCGAGCCT	677
QY	679	TGCGCACCAAGTGTGCGGCGTCCATGAAGCGAAAAAGAAATTGCTAGTCAACCCGAATAA	738
Db	678	TGCGCACCAAGTGTGTGATCGACATGAGGCGAAGAAAAATTCTGTTAACATCTTGAGCAA	737
QY	739	AATAATCATGCAA 751	
Db	738	AATAATTAATCTAA 750	

```

RESULT 5
CB038112
LOCUS
DEFINITION
CB038112 766 bp mRNA linear EST 15-JAN-2003
TC ad2_42A03_TEXF1 Teladorsagia circumcincta adults library 2
Teladorsagia circumcincta cDNA clone TC ad2_42A03 5' similar to
CAD56659 Ancylostoma-secreted protein-like prote Ostertagia, mRNA
sequence.
ACCESSION
CB038112
VERSION
CB038112.1 GI:27757357
KEYWORDS
EST.
SOURCE
Teladorsagia circumcincta
Teladorsagia circumcincta
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
Trichostrongyloidea; Haemonchidae; Ostertagiinae; Teladorsagia.
REFERENCE
1 (bases 1 to 766)
Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D.,
Hall,N., Quayle,M. and Barrell,B.
Edinburgh University/Sanger Centre Nematode EST Project
Unpublished (2000)
COMMENT
Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared by Dr Diane Redmond and Dr David Knox,
Moredun Research Institute, Midlothian, UK.
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.
Sequencing was performed by the Pathogen Sequencing Unit, Sanger
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).
PCR PRIMERs
FORWARD: TEXPCR.F1
BACKWARD: T7.P1
Plate: 42 row: A column: 03
Seq primer: TEXF1
High quality sequence stop: 516.
location/Qualifiers
1. .766
/mol_type="mRNA"
/organism="Teladorsagia circumcincta"

```



/db\_xref="taxon:45464"  
/clone="Tc ad2\_42A03"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site\_1: BamHI; Site\_2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match 62.1%; Score 474.2; DB 6; Length 766;  
Best Local Similarity 79.0%; Pred. No. 2.1e-132;  
Matches 579; Conservative 0; Mismatches 148; Indels 6; Gaps 1;  
QY 19 GCGTGTGTAGTGTCTGTTCTCTGCGCCCTGTTCTCTATGCGGAAGCAGGCTTTG 78  
Db 25 GGCAGTCACAGTGTCTGTTCTCTGCGCTTGTCTCTCCATGCCAGGCGAGCTTTG 84  
QY 79 TCCGAATAGTCTAAGCCAAAGTGAAGCGGAGGAGGAGATTTCTCGATTTCAATGA 138  
Db 85 CCCAAAACCTTAAGCCAAAGTGAAGCGGAGGAGGAGATTTCTCGATTTCAATGA 144  
QY 139 TGTTCGTGAATATAGCACTTGGAATGTTGATTAATCTGACAGTAATGCAAGCC 198  
Db 145 TATTCTCGGAACATAGCACTTGGAAGAGCTTGTGAATCT-----TCAGACTGATCC 198  
QY 199 GGTGATTTCTGTCAGCTCAGAACATGTACAAAGTGACTGGGATTCGAATTTGAAGA 258  
Db 199 GGTGATTTCTGTCAGCTCAGAACATGTACAAAGTGACTGGGATTCGAATTTGAAGA 258  
QY 259 AGTAGCAGACACACAGATGCGGCATGCAATGATCCCTTACCGATAATACCAAGCTGCG 318  
Db 259 AAAAGCAGCTCAACAGATGCGGCATGCAAGCTGCTCTTACCAATAGATCCCAAGCTGCG 318  
QY 319 TCAAAATATCGTAGATGCGTGTACTTCAAGACAGTAAGAGAAGACAGTTCTGACAA 378  
Db 319 TCAAGATATCGTAGATGCGTGTACTTCAAGACAGTAAGAGAAGACAGTTCTGACAA 378  
QY 379 AGTATCGTGTGTTGGGTGAGGCAATCGCTGGGATTTATGAAGGACGAACTTGACCA 438  
Db 379 AGTATCGTGTGTTGGGTGAGGCAATCGCTGGGATTTATGAAGGACGAACTTGACCA 438  
QY 439 ATTGCTAACCAAGTGGGCTGAACCTTAGCAAAACATGCAAACTATAGAAACGAAAGGT 498  
Db 439 TTTTGCTAACCAAGTGGGCTGAACCGCTTGCAACATGCAAACTGAAAAATCGGAAGGT 498  
QY 499 TGGATGTGCCCATAGATCTGCCCCGCTCAGCAAAACAGTAGTATCTGCTGCTATGG 558  
Db 499 TGGATGTGCCCATAGATCTGCCCCGCTCAGCAAAACAGTAGTATCTGCTGCTATGG 558  
QY 559 AAGCCCCAAACTTGCAAGCAAGAAATATATCTGCGAAGAAAGGCTTGTGTGCGA 618  
Db 559 AAGCCCCAAACTTGCAAGCAAGAAATATATCTGCGAAGAAAGGCTTGTGTGCGA 618  
QY 619 CGCTCGTCCAGATTCATCTGCTGCGAACAACCTGTGTGACACGCGAGATGCTGCGAGTGT 678  
Db 619 CTCTTATCCGATTCGTTTGTGCTGTGACAGCCTGTGTGACACGAGATGCTGCGAGTGT 678  
QY 679 TCGCCACCAAGTGTGCGGCTGCGCATGAAAGCGAAAGAAATGTTAGTCAAGCCGCAATGA 738  
Db 679 TCGCCACCAAGTGTGCGGCTGCGCATGAAAGCGAAAGAAATGTTAGTCAAGCCGCAATGA 738  
QY 739 AATATTCATGCAA 751  
Db 739 AATATTCATGCAA 751

RESULT 6  
CB036391 756 bp mRNA linear EST 15-JAN-2003  
LOCUS Tc ad2\_21B12\_TEXF1 Teladorsagia circumcincta adults library 2  
DEFINITION Teladorsagia circumcincta cDNA clone Tc ad2\_21B12 5' similar to

CAD56659 Ancylostoma-secreted protein-like prote Osterlagia, mRNA  
sequence.  
CB036391  
CB036391.1 GI:27755636  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guillian, D.,  
Hall, N., Quayle, M. and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
COMMENT  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JF, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCR1  
BACKWARD: T7PL  
Plate: 21 row: B column: 12  
Seq primer: TEXF1  
High quality sequence stop: 518.  
Location/Qualifiers

FEATURES

Source

1. 756  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="Tc ad2\_21B12"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site\_1: BamHI; Site\_2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match 62.1%; Score 473.6; DB 6; Length 756;  
Best Local Similarity 79.3%; Pred. No. 3.2e-132;  
Matches 576; Conservative 0; Mismatches 144; Indels 6; Gaps 1;  
QY 19 GCGTGTGTAGTGTCTGTTCTCTGCGCCCTGTTCTCTATGCGGAAGCAGGCTTTG 78  
Db 25 GGCAGTCACAGTGTCTGTTCTCTGCGCTTGTCTCTCCATGCCAGGAGGCTTTG 84  
QY 79 TCCGAATAGTCTAAGCCAAAGTGAAGCGGAGGAGGAGATTTCTCGATTTCAATGA 138  
Db 85 CCCAAAACCTTAAGCCAAAGTGAAGCGGAGGAGGAGATTTCTCGATTTCAATGA 144  
QY 139 TGTTCGTGAATATAGCACTTGGAATGTTTGAATACTGACAGTAATGCAAGCC 198  
Db 145 TATTCTGGAACAGACACTTGGAAGAGCTTGTGAATCT-----TCAGACTGATCC 198  
QY 199 GGTGATTTCTGTCAGCTCAGAACATGTACAAAGTGACTGGGATTCGAATTTGAAGA 258  
Db 199 GGTGATTTCTGTCAGCTCAGAACATGTACAAAGTGACTGGGATTCGAATTTGAAGA 258  
QY 259 AGTAGCAGACACACAGATGCGGCATGCAATGATCCCTTACCGATAATACCAAGCTGCG 318  
Db 259 AAAAGCAGCTCAACAGATGCGGCATGCAAGCTGCTCTTACCAATAGATCCCAAGCTGCG 318  
QY 319 TCAAAATATCGTAGATGCGTGTACTTCAAGACAGTAAGAGAAGACAGTTCTGCAACA 378

||||| 319 TCAAAATATCGCAAGATGGCTGTATTATGCCAACATGAGAAACAAGGTTTGAGACA 378  
QY 379 AGTATCGTGTATGGGTGAGCGCATCGCTGGATTATGAAAGGACGAAACTTGACCA 438  
DB 379 AGTACCGTGTCTGGGTAAACCCCATCGCTACGATTTATGAAAGGCGACGGCACTTGATCG 438  
QY 439 ATTTGCTAACCACTGGCGCTGAAACCTCTAGCAAACTTGCAAACTATAGAAAACGAAAGGT 498  
DB 439 TTTTGCTAACCACTGGCGCTGAAACCGCTTGCCAACTGTCTAACTGAAAAATCGAAAGGT 498  
QY 499 TGGATGTGCCCATAGATCTGCCCGCTCAGCAAAACATGAGTAGTATCCTGCGTGTATGG 558  
DB 499 TGGATGTGCTTACAAATCTGCCCCCGCTGAAAAACATGTCGTATCCTGTGTATGG 558  
QY 559 AAGCCCCCAACTTGCACCGCAAGAAAGTATCTGGCAGGAAGAAAGGCTTGTTGTGCGA 618  
DB 559 AAGCCAAAACCTTTCACCAACGAAAGTAATTTGGGAGCAAGGAAGTACTTGCAGTGCAA 618  
QY 619 CGCTCGTCCAGATTCACTCTGCTGCGACCAACCTGTGTGACACGCGAGATGCTGCGAGTGT 678  
DB 619 CTCTTATCCGATTGTTTGTGCTGTGACAGCCTGTGTGACACGCGATGAGCTGCGAGCCT 678  
QY 679 TCGCCACCAAGTGTGCGCGTGCCTCATGAAAGCAAAATTGGTAGTCAACCCCGAATAA 738  
DB 679 TCGCCACCAAGTGTGTTGATGCAATGAGGCGAAGAAAATTCTGTTAACATCTTGAGCAA 738  
QY 739 AATATT 744  
DB 739 AATATT 744

RESULT 7  
CB036975 755 bp mRNA linear EST 15-JAN-2003  
LOCUS  
DEFINITION TC ad2\_28D02\_TEXF1 Teladorsagia circumcincta adults library 2  
Teladorsagia circumcincta cDNA clone TC ad2\_28D02 5' similar to  
CAD56659 Ancylostoma-secreted protein-like prote Osteragia, mRNA  
sequence.  
ACCESSION CB036975  
VERSION CB036975.1 GI:27756220  
KEYWORDS EST.  
SOURCE Teladorsagia circumcincta  
ORGANISM Teladorsagia circumcincta  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Strongylida;  
Trichostrongyloidea; Haemonchidae; Ostertagiinae; Teladorsagia.  
REFERENCE 1 (bases 1 to 755)  
AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Gulliano,D.,  
Hall,N., Quayle,M. and Barrell,B.  
TITLE Edinburgh University/Sanger Centre Nematode EST Project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCRFL  
BACKWARD: T7PL  
Plate: 28 row: D column: 02  
Seq primer: TEXF1  
High quality sequence stop: 516.

FEATURES  
Source 1..755  
/organism="Teladorsagia circumcincta"

/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="TC\_ad2\_28D02"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site\_1: BamHI; Site\_2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN  
Query Match 62.0%; Score 473.2; DB 6; Length 755;  
Best local Similarity 79.3%; Pred. No. 4.2e-132;  
Matches 575; Conservative 0; Mismatches 144; Indels 6; Gaps 1;

QY 20 GCTGTGTAGTTGCTGTTCTCTCGCCCTGTTCTCCTATGCGGAGCAGGCTTTGTTGT 79  
DB 27 GCTGTACAGTTGTTGCTCTCTCGCTCTGCTCTGCTCTATGCGGAGCAGGCTTTGTTGC 86  
QY 80 CCGAATAGTCTTAAGCCAAAGTGACAGCGCGGAGCAGATTTCCTCGATTTCACATGAT 139  
DB 87 CCAAAACTCTGGGCCAAACCGACCAACGCAAGCAGATTTCCTCGATTTCACATGAT 146  
QY 140 GTTGTGCAAAATATAGCACTTGAAATGTTGATTAACCTGACAGTAAATGCAAGCGG 199  
DB 147 ATTGTGGAACATAGCACTTGAAAGACTTGTTGTAAGTCTGTAAGTCTGTAAGTCTG 200  
QY 200 GTCATTTCTGTCAGCTCAGAACATGTACAAAGTGAAGTGGATTGCACTTGGAAGAA 259  
DB 201 GTTATTTCTGTCAGCTCAAAACATGTACAGCTGATTGGAGTGTGACTTGGAACAA 260  
QY 260 GTAGCAGCAACAGATGGCGCCATGCAATGATCCCTACCGATAAATACCAAGCTGGCT 319  
DB 261 AAGCAGCCAGCAGATGGCGCAATGACAGGTTCTCTACCAATAGATCCAGCTGGCT 320  
QY 320 CAAATATCGCTAGATGCTGTACTTCAAAGACATGAAAGACAGTCTTGCAACAA 379  
DB 321 CAAATATCGCAAGATGGCTGTATTATGCCAACAGTGAAGAACAGGTTTGAGACAA 380  
QY 380 GTATCGTGTATTGGGTGAGCGCATCGCTGGATTATTAAGAAAGCAGAACTTGACCAA 439  
DB 381 GTACCGTGTCTTGGGTAAACCCCATCGCTACGATTTATGAAGGGCACGGCACTTGATCGT 440  
QY 440 TTGCTAACCACTGGGTGAACCTCTAGCAAACTATAGAACTATAGAAACCGAAAGGT 499  
DB 441 TTGCTAACCACTGGGTGAACCGCTTGCCAACTGCTTAAGTGAAGAAATCGGAAGTT 500  
QY 500 GGATGTGCCCATAGATCTGCCCGCTCAGCAAAACATGAGTATCCTGCTGATGGA 559  
DB 501 GGATGTGCCCATAGATCTGCCCGCTCAGCAAAACATGAGTATCCTGCTGATGGA 560  
QY 560 AGCCCCCAACTTGCACCGCAAGAAAGTATCTGCGAGGAAGAAAGGCTTGTTGTGCGAC 619  
DB 561 AGCCAAAACCTTTCACCAACGAAAGTAATCTGGGAGCAAGAAATACCTTGCAAGTGAAC 620  
QY 620 GCTGTCCAGATTCACTCTGCTGCGACCAACCTGTGTGACACGCGAGATGCTGCGAGTGT 679  
DB 621 TCTTATCCGATTGTTTGTGCTGTGACAGCCTGTGTGACACGCGATGCTGCGAGCCTT 680  
QY 680 CGCCACCAAGTGTGCGCGTGCCTCATGAAAGCAAGAAATGTTGTTGTTGTTGTTGTT 739  
DB 681 CGCCACCAAGTGTGTTGATGCAATGAGGCGAAGAAATTTCTGTTAATCTTGAGCAAA 740  
QY 740 AATATT 744  
DB 741 AATATT 745

RESULT 8  
CB038323 751 bp mRNA linear EST 15-JAN-2003  
LOCUS  
DEFINITION TC\_ad2\_44D11\_TEXF1 Teladorsagia circumcincta adults library 2





/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="Tc\_ad2\_43A10"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site 1: BamHI; Site 2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match	61.9%;	Score 472.6;	DB 6;	length 754;
Best Local Similarity	78.9%;	Pred. No. 6.4e-132;		
Matches 578;	Conservative 0;	Mismatches 149;	Indels 6;	Gaps 1;
OY	19	GGCTGTTGAGTGTGCTGTTCTCTGGCCCTGTTCTCTATGCCGAAGCAGGCTTTGTG	78	
Db	25	GGCAGTCACAGTTGCTGTTCTTCTGGCTTGTCTCCATGCCGAGGCGCTTTGTG	84	
OY	79	TCCGAATAGTCTAAGCCAAAGTGAAGCGCGAGCAGATTTTCTCGATTTTCACAATGA	138	
Db	85	CCCAAAAACCTTAAGCCAACTGACAACGCGAGCAGATTTTCTCGATTTACACATGA	144	
OY	139	TGTTGTCGAATATAGCACTTGAATGTTGATTAATCTGACAGTAATGCAAGCGC	198	
Db	145	TATTCGTGGAACATAGCACTTGAAGAAGCTTGTGAAGT-----TCACGACTGATCC	198	
OY	199	GGTCATTTCTGTGCCAGCTCAGAACATGTACAAGTGAAGTGGGATTTGCAACTTGAAGA	258	
Db	199	GGTTGTTCTTGTCAGCTCAGAACATGTACAAGTGAAGTGGGATTTGCAACTTGAAGA	258	
OY	259	AGTAGACACACAACAGATTGCCCATGCAATGATCCCCCTAACCGATAATACCAGCTGGC	318	
Db	259	AAAAGCAGCTCAACAGATTGCCCATGCAAGGTTCTCTACCAATAGATCCCAAGCTGGC	318	
OY	319	TCAAAATATCGCTAGATGGCTGTACTTCAAAAGCAGTGAAGAAGACAGTCTTGACACA	378	
Db	319	TCAAAATATCGCAAGATGGCTGTATATGCCAACAGTGAAGAAGACAAAGTTTGAGACA	378	
OY	379	AGTATCGTGATTTGGGTGAGCGCATCGCTGGGATTTATGAAGGACGAAACTTGACCA	438	
Db	379	AGTACCGTGGTCTTGGGTAAACCCGCTACGATTCAATGAAGGACACGCACTTGATCG	438	
OY	439	ATTGCTAACCAAGTGGGCTGAACCTTAGCAAAACATTGCAAACTATAGAAACCGAAAGGT	498	
Db	439	TTTTGCTAACCAAGTGGGCTGAACCGCTTGCCCAACATTGCTAAGTGAAGAAATCGAAGGT	498	
OY	499	TGGATGTGCCCATTAAGATCTGCCCGCTCAGCAAAACATGCTATCTCTGCTGATGG	558	
Db	499	TGGATGTGCTTAACAAGATCTGCCCGCTGAAAAACATGCTGCTATCTGTGTATGG	558	
OY	559	AAGCCCCCAACTTGCAACCGAAGGTAATCTGGCAGGAAGGAAGGCTGTGTGCGA	618	
Db	559	AAGCCAAAACCTTCAACCAACGAAGTAATTTGGAGCAAGGAATTAATTGCGAGTGCAA	618	
OY	619	CGCTGTCAGATTCATTCTGCTGCGACAACCTGTGTGACACGCGAGATGCTGCGAGTGT	678	
Db	619	CTCTATCCGGAATCGTTTGTGCTGACAGCTGTGTGACACGAATGAGCTGCGAGCT	678	
OY	679	TCCGACCAAGTGTGCGCGCTGCCATGAAGCAAAAGAAATTTGTAAGTCAACCCGAATAA	738	
Db	679	TCCGACCAAGTGTGTGATGCACATGAGGGAAGAAATTTCTGTTAACATCTTGAGCAA	738	
OY	739	AATATTCAATGCAA 751		
Db	739	AATATTATTCTAA 751		

RESULT 10

CB038345

LOCUS CB038345 771 bp mRNA linear EST 15-JAN-2003

DEFINITION Tc\_ad2\_44F10 TEXF1 Teladorsagia circumcincta adults library 2  
Teladorsagia circumcincta cDNA clone Tc\_ad2\_44F10 5' similar to  
CAD56659 Ancylostoma-secreted protein-like prote Ostertagia, mRNA  
sequence.  
ACCESSION CB038345  
VERSION CB038345.1 GI:27757590  
KEYWORDS EST.  
SOURCE Teladorsagia circumcincta  
ORGANISM Teladorsagia circumcincta  
Trichostrongylidae; Haemonchidae; Ostertagiinae; Teladorsagia.  
1 (bases 1 to 771)  
Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guillian,D.,  
Hall,N., Quayle,M. and Barrell,B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
COMMENT Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCR1  
BACKWARD: T7PL  
plate: 44 row: F column: 10  
Seq primer: TEXF1  
High quality sequence stop: 517.  
location/Qualifiers  
1..771  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="Tc\_ad2\_44F10"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site 1: BamHI; Site 2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match	61.6%;	Score 470.2;	DB 6;	length 771;
Best Local Similarity	78.9%;	Pred. No. 3.4e-131;		
Matches 575;	Conservative 0;	Mismatches 148;	Indels 6;	Gaps 1;
OY	19	GGCTGTTGAGTGTGCTGTTCTCTGGCCCTGTTCTCTATGCCGAAGCAGGCTTTGTG	78	
Db	24	GGCAGTCACAGTTGTTCTCTGCTTGTGTCTCCATGCCGAAGCAGGCTTTGTG	83	
OY	79	TCCGAATAGTCTAAGCCAAAGTGAAGCGCGAGGCAAGATTTTCTCGATTTTCACAATGA	138	
Db	84	CCCAAAAACCTTAAGCCAACTGACAACACTAGGACAGATTTTCTCGATTTTCACAATGA	143	
OY	139	TGTTGTCGAATATAGCACTTGAATGTTGATTAATCTGACAGTAATGCAAGCGC	198	
Db	144	TATTCGTGGAACATAGCACTTGAAGAAGCTTGTGAAGT-----TCACGACTGATCC	197	
OY	199	GGTCATTTCTGTGCCAGCTCAGAACATGTACAAGTGAAGTGGGATTTGCAACTTGAAGA	258	
Db	198	GGTTGTTCTTGTCAGCTCAGAACATGTACAAGCTGGAAGTGGGATTTGCAACTTGAAGA	257	
OY	259	AGTAGCAGACACAACAGATTGCCCATGCAATGATCCCCCTAACCGATAATACCAGCTGGC	318	
Db	258	AAAAGCAGCTCAACAGATTGCCCATGCAAGGTTCTCTACCAATAGATCCCAAGCTGGC	317	

OY		319	TCAAAATATCGCTAGATGGCTGTACTTCAAAGA CAGTGAGGAAGACAGTTCTGCAACA	378
Db		318	TCAAAATATCGCAAGATGGCTGTATTATTCCAA CAGTGAGGAGACAAGTTTTGAGACA	377
OY		379	AGTATCGTGTATTGGGTGAGCCGATCGCTGGG ATTTATGMAAGGCAGAAA CTTGACCA	438
Db		378	AGTACC GTGCTTGGGTAA CCCCCTACGCTA CGATTCATGMAAGGCCACGGC ACTTGATCG	437
OY		439	ATTGCTAAC CAGTGGGCTGAACCTTAGCAAA CATTGCAAACTATA GAAACGAAAGGT	498
Db		438	TTTCGCTAAC CAGTGGGCTGAACCGCTTGCCA ACATTGCTAACTGMAAAATCGGAAGT	497
OY		499	TGGATGTGCCCATTA GATCTGCCCGCTCAGCA AAACATGTA GTATCTGCGTGTATGG	558
Db		498	TGGATGTGCTTACA GATCTGCCCGCCTGAAAA ACATGTCGTATCTGTGTATGG	557
OY		559	AAGCCCCAA ACTTGCA CCGGAAGATTA TCTGGCAGGAAGAA AGGCTGTGTGTGCGA	618
Db		558	AAGCCAAA AAC TTTCACCA CAAGATTA ATTGGAGCAAGSAA GTA CTTCGAGTGCAA	617
OY		619	CGCTCGTCCAGAT TCATTCTGTGCCA CAACCTGTGTGACA CCGGAGATGCTGCCAGTGT	678
Db		618	CTCTTAATCCGGA ATCGTTTGTGTGTGACA CAGCCTGTGTGACA CCGCGTGAAGCTGCCAGCCT	677
OY		679	TCGCCAACCA GTGTGCGCGCTGCCCATGA AGCGAAATTTGTA GTCA CCCCCGAATTA	738
Db		678	TCGCCAACCA GTGTGCTGATCGACATGA GGCGGAAGAATTCTGT TAA CATTGTGAGCAA	737
OY		739	AATATTCA T 747	
Db		738	AATATTCA T 746	

[illegible]

```

source

1. .733
/organism="Teladorsagia circumcincta"
/mol_type="mRNA"
/db_xref="taxon:45464"
/clone="TC_ad2_36B07"
/sex="mixed"
/dev_stage="adult"
/clone_1lb="Teladorsagia circumcincta adults library 2"
/note="Vector: pTIP1bX2; Site_1: BamHI; Site_2: BamHI;
Teladorsagia circumcincta is a parasitic nematode. The
library was constructed from mRNA from Teladorsagia
circumcincta mixed adults."

```

Query Match	61.5%	Score 469.2	DB 6	Length 733
Best Local Similarity	79.8%	Pred. No. 6.8e-131		
Matches 568	Conservative 0	Mismatches 138	Indels 6	Gaps 1
QY	19	GGCTGTGTAGTGTCTGTCTCTCGGCCCTGTCTCTATGCCGAAGCAGGCTTTGTTG	78	
DB	25	GGCAGTCACAGTTGCTGTCTCTCTGTGCTTGTCTCCATGCCGAAGCAGGCTTTGTTG	84	
QY	79	TCCGAATAGTCTAAGCCAAAGTGACAGCGCGAGGAGATTTCTCGATTTTCACAATGA	138	
DB	85	CCCAAAACTCTAAGCCAAACTGACACGCGAGGAGATTTCTCGATTTTCACAATGA	144	
QY	139	TGTTGTCGAAATATAGCACTTGGAATGTTGATTAACCTGACAGTAATGCAGCGC	198	
DB	145	TATTGTCGAAACATAGCACTTGGAAGAGCTTGTGTGAAGTCTTCTGATGATCC	198	
QY	199	GGTCATTCTTGGTCCAGCTCAGAACATGTACAAAGTGAACTGGGATTGCCACTTGAA	258	
DB	199	GGTTGTTCTTGGTCCAGCTCAGAACATGTACAAAGTGAACTGGGACTGTGACCTGAA	258	
QY	259	AGTAGCAGCACACAGATTGCGCCATGCATGATCCCTACCGATAAATACAGCTGGC	318	
DB	259	AAAGCAGCTCAACAGATTGCGCAATGACAGGTTCTCTACCGATAGATCCAGTTGGC	318	
QY	319	TCAAATATCGCTAGATGCGCTGTACTTCAAAGCAGTGAAGAAGACAGTTCTGACA	378	
DB	319	TCAAATATCGCAGATGCGCTGTATATGCCAAAGTGAAGAAGACAGGTTTGAGACA	378	
QY	379	AGTATCGTGTATGGGTGAGCGCATCGCTGGATTATGAAGGACGAAACTTGACCA	438	
DB	379	AGTACCGTGTCTTGGGTAAACCCCATCGCTACGATTCATGAAGGACGCGCACTGATCG	438	
QY	439	ATTGCTAACCAAGTGGGCTGAACCTCTAGCAACATTTGCCAACTATAGAAACCGAAAG	498	
DB	439	TTTGTCTAACCAAGTGGGCTGAACCGCTTGCCAACATTTGCTAACCTGAAAAATCGGAAG	498	
QY	499	TGATGTGCCCATAGATCTGCCCCGCTCAGCAAAACATGCTATCTCTGCTGTATGG	558	
DB	499	TGATGTGCTTACAGATCTGCCCGCCCTGAAAAACATGCTGTATCTGTGTATGG	558	
QY	559	AAGCCCCAACTTGCAACCGAAGATTATCTGCGAAGAAAGGCTTGTGTGCGA	618	
DB	559	AAGCCAAAACTTCAACCAACGAAGTATTTGGAGCAAGGAAGTACTTGCGAGTGCAA	618	
QY	619	CGCTGTCCAGATTCATTCGTCTGCGACCAACCTGTGTGACACGCGAGATGCTGCGAG	678	
DB	619	CTCTTATCCGGAATTCGTTTGTCTGTGACAGCCCTGTGTGACACGCGATGAGCTGCGAG	678	
QY	679	TGCGCACCAAGTGTGCGCGCTGCGCCATGAAAGCAAAAGAAATTGTGATGACACC	730	
DB	679	TGCGCACCAAGTGTGCTGTGATGACATGAGGCGGAGAAAAATTCTGTTAACATC	730	

RESULT 12

LOCUS CB037329

DEFINITION CB037329 763 bp mRNA linear EST 15-JAN-2003

Tc ad2\_32F03 TEXF1 Teladorsagia circumcincta adults library 2

Teladorsagia circumcincta cDNA clone Tc ad2\_32F03 5' similar to

CAD5659 Ancylostoma-secreted protein-like prote Osetetragia, mRNA

61.5% Score 469.2; DB 6; Length 733;

Best Local Similarity 79.8%; Pred. No. 6.8e-131;

Matches 568; Conservative 0; Mismatches 138; Indels 6; Gaps 1;

ACCESSION	sequence.
VERSION	CB037329
KEYWORDS	CB037329.1 GI:27756574
SOURCE	EST.
ORGANISM	Teladorsagia circumcincta Teladorsagia circumcincta Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida; Trichostrongylidae; Haemonchidae; Ostertagiinae; Teladorsagia. 1 (bases 1 to 763)
REFERENCE	Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D., Hall,N., Quayle,M. and Barrell,B. Edinburgh University/Sanger Centre Nematode EST Project Unpublished (2000)
TITLE	Contact: Blaxter ML
JOURNAL	Institute of Cell, Animal and Population Biology University of Edinburgh
COMMENT	Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK. Tel: +44 131 650 6760 Fax: +44 131 670 5450 Email: mark.blaxter@ed.ac.uk The library was prepared by Dr Diane Redmond and Dr David Knox, Moredun Research Institute, Midlothian, UK. PCR was performed by Ye Jieru, ICAPB, University of Edinburgh. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). PCR Primers FORWARD: TEXPCRF1 BACKWARD: T7PL Plate: 32 row: F column: 03 Seq primer: TEXF1 High quality sequence start: 20 High quality sequence stop: 549. Location/Qualifiers 1..763 /organism="Teladorsagia circumcincta" /mol_type="mRNA" /db_xref="taxon:45464" /clone="Tc_ad2_32F03" /sex="mixed" /dev_stage="adult" /clone_lib="Teladorsagia circumcincta adults library 2" /note="Vector: pTriplEx2; Site_1: BamHI; Site_2: BamHI; Teladorsagia circumcincta is a parasitic nematode. The library was constructed from mRNA from Teladorsagia circumcincta mixed adults."
FEATURES	
source	
ORIGIN	
Query Match	61.5%; Score 469.2; DB 6; Length 763;
Best Local Similarity	79.8%; Pred. No. 6.9e-131;
Matches 568; Conservative	0; Mismatches 138; Indels 6; Gaps 1
OY	19 GGCTGTTGAAGTGTCTGTCTTCCTGGCCCTGTTCTCCTATGCCGAGCAGCGTTTGTTG 78
Db	55 GGCAGTCAAGTTGCTGTTCTTCTTGCTTTGTCTCCATGCCGAGCAGCGTTTGTTG 114
OY	79 TCCGAATAGTCTAAGCCAAAGTGACAGCGGAGGAGCATTTTCTCGATTTCACAATGA 138
Db	115 CCCAAAACCTTAAGCCCAAAGTGACCAACGCGAGGAGCATTTTCTCGATTTCACAATGA 174
OY	139 TGTTGTCGAATAATAGCACTTGAAATGTTTGATAACTGACAGTAATGCAGACGC 198
Db	175 TATTGTCGAACATAGCACTTGAAAGAGCTTGTAAGT-----TCAGACTGATCC 228
OY	199 GGTCAATCTTGTTCCAGCTCAGAACATGTACAAAGTGGACTGGGATTGCAACTGGAAGA 258
Db	229 GGTGTTCTTGTTCCAGCTCAGAACATGTACAAAGTGGACTGGGACTGTGACCTGGAACA 288
OY	259 AGTAGCAGCACACAGATTGCGCCATGCATGATCCCCTACCAGTAATACAGCCTGGC 318
Db	289 AAAAGCAGCTCAACAGATTGCGCAATGCAGGGTTCCTTACCGATAGATCCAGTCTGGC 348
OY	319 TCAAAATATCGCTAGATGGCTGTACTTCAAGAAGAGTGAAGAAGAGACAGTTCTGCAACA 378

	Db	349	TCAAAATATCCGAAGATGGCTGTATATGCCAACAGTGAAGAACAAGGTTTGAGACA	408
Oy	379	AGTATCGTGGTAATTGGGTGAGCGCATCGCTGGGATTATGAAAGCACGAACCTTGACCA	438	
Db	409	AGTACC GTGTC TTGGGTAA CCCC CATCGCTA CGATT CATGAA GGGCAC GGCA CTGATCG	468	
Oy	439	ATTGCTAACAGTGGGCTGAACCTTAGCAAACATTGCMAACTATAGAAAACCGAAAGT	498	
Db	469	TTTTGCTAACCA GTGGGCTGA ACCGCTTGCCA ACATTGCT TAACTGGA AAAATCGGAAAGT	528	
Oy	499	TGGATGTGCCATAAGATCTGCCCCGCTCAGCAAAACATGTTAGTATCTGCGTGTATGG	558	
Db	529	TGGATGTGCTTACAAGATCTGCCCGCCCTGAAAAACATGTCGTATCTGTGTATGG	588	
Oy	559	AAGCCCAAACTTGCA CCGAACGA AGTTATCTG GCAGAGAAAGC TTGTGTGCGCA	618	
Db	589	AAGCCAAAACTTTCACCAACGAAGTATTTGGAGCAAGAAAGTACTTGCGAGTGCAA	648	
Oy	619	CGCTCGTCCAGATTCA TTCCTGCTGCGCA CAACCTGTGTGACACCGCAGATGCTCGAGTGT	678	
Db	649	CTCTTATCCGATTGTTGTTGTGCTGTGACAGCCTGTGTGACACGCATGAGCTGCGAGCCT	708	
Oy	679	TCGCCACCAAGTGTGCGCGCTGCCCATGAAGCGAAAGAAATTGGTAGTCACC	730	
Db	709	TCGCCACCAAGTGTGTTGATGCATGAGCGCAGAAAAATTCTGTAAACATC	760	
<b>RESULT 13</b>				
CB036489		739 bp	mRNA	linear EST 15-JAN-2003
LOCUS				
DEFINITION	Cb036489 CB036489.1 GI:27755734			
VERSION	Teladorsagia circumcincta adult library 2			
KEYWORDS	EST.			
SOURCE	Teladorsagia circumcincta			
ORGANISM	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;			
	Trichostrongylidae; Haemonchidae; Oostertaginae; Teladorsagia.			
	1 (bases 1 to 739)			
REFERENCE	Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D.,			
AUTHORS	Hall,N., Quayle,M. and Barrell,B.			
TITLE	Edinburgh University/Sanger Centre Nematode EST Project			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Blaxter ML Institute of Cell, Animal and Population Biology University of Edinburgh Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK. Tel: +44 131 650 6760 Fax: +44 131 670 5450 Email: mark.blaxter@ed.ac.uk The library was prepared by Dr Diane Redmond and Dr David Knox, Moredun Research Institute, Midlothian, UK. PCR was performed by Ye Jieru, ICAPB, University of Edinburgh. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).			
<b>FEATURES</b>				
source	1..739 location/Qualifiers organism="Teladorsagia circumcincta" mol_type="mRNA" db_xref="taxon:45464" clone="Tc_ad2_22E03" sex="mixed"			



/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site\_1: BamHI; Site\_2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match 61.3%; Score 467.6; DB 6; Length 739;  
Best Local Similarity 79.6%; Pred. No. 2.1e-130;  
Matches 567; Conservative 0; Mismatches 139; Indels 6; Gaps 1;

QY 19 GGCTGTGTAGTGTCTTCTCTGCGCCCTGTTCTCCTATGCGGAGCAGCCTTTGTTG 78  
DB 31 GGAGTCAAGTGTCTGTTCTTCTGCTTGTCTCCATGCGGAGCAGCCTTTGTTG 90  
QY 79 TCCGAATAGTCTAAGCCCAAGTGAACGCGGAGCAGATTTCTCGATTTTCACAATGA 138  
DB 91 CCCAAAACCTTAAGCCCAAACTGACACGCGAGCAGATTTCTCGATTTTCACAATGA 150  
QY 139 TGTTCGTGAAATATAGCACTTGAAATGTTGATTAACCTGACAGTAATGACAGCGC 198  
DB 151 TATTGTCGGAACATAGCACTTGAAAGCTTGTAAGT-----TCACGACTGATCC 204  
QY 199 GGTCAATCTTGTCCAGCTCAGAACATGTACAAAGTGAAGTGGAGTGAAGTGAAGA 258  
DB 205 GGTGTTCTTGTCTCAGCTCAGAACATGTACAAAGTGAAGTGGAGTGAAGTGAAGA 264  
QY 259 AGTAGCAGCACAACAGATTGGCCCATGCAATGATCCCTACCGATTAATACCAAGCTGGC 318  
DB 265 AAAAGCAGCTCAACAGATTGGCCCATGCAATGATCCCTACCGATTAATACCAAGCTGGC 324  
QY 319 TCAAAATATCGTAGATGGCTGATCTTCAAAAGCAGTGAAGAGAGACAGTTCTGACACA 378  
DB 325 TCAAAATATCGTAGATGGCTGATCTTCAAAAGCAGTGAAGAGAGACAGTTCTGACACA 384  
QY 379 AGTATCGTGTATTTGGGTGAGCGCATCGCTGGGATTTATGAAGCAGCAAACTTGACCA 438  
DB 385 AGTACCGTGTCTTGGGTGAGCGCATCGCTGGGATTTATGAAGCAGCAAACTTGATCG 444  
QY 439 ATTGCTAACAGTGGGCTGAACCTCTAGCAACATTGCAAACTATAGAAACCGAAGGT 498  
DB 445 TTTTGTAAACAGTGGGCTGAACCTCTAGCAACATTGCTAACTGCAAAATCGGAAGGT 504  
QY 499 TGGATGTGCCATTAAGATCTGCCCCGCTCAGCAAAACATGTGATCTCTGCGTATGG 558  
DB 505 TGGATGTGCTTAACAAGATCTGCCCCGCTCAGCAAAACATGTGATCTCTGCGTATGG 564  
QY 559 AAGCCCCAAACTTGCAACCGAAGTATCTGGCAGGAGGAAAGGCTGTGTGCGCA 618  
DB 565 AAGCCAAAACCTTTCACCAACGAAGTATTTGGAGCAAGGAATTAATTGCGAGTGCAA 624  
QY 619 CGCTGCTCAGATTCATTCTGCTGCGCAACCTGTGTGACACGCGAGATGCTGCGAGTGT 678  
DB 625 CTCTTATCCGATTCGTTTGTGCTGACAGCTGTGTGACACGAATGAGCTGCGAGCTT 684  
QY 679 TCGCCACAGTGTGCGCGCTGCGCATGAGCGAAAGAAATTTGTTAGTCAAC 730  
DB 685 TCGCCACAGTGTGTTGATGACATGAGCGGAGAAATTTCTGTTAATC 736

RESULT 14

LOCUS CB037895 800 bp mRNA linear EST 15-JAN-2003  
DEFINITION TC\_ad2\_39D08\_TBxP1 Teladorsagia circumcincta adults library 2  
Teladorsagia circumcincta cDNA clone TC\_ad2\_39D08 5' similar to  
CAD56659 Ancylostoma-secreted protein-like prote Ostertagia, mRNA  
sequence.  
ACCESSION CB037895  
VERSION CB037895.1 GI:27757140  
KEYWORDS EST.  
SOURCE Teladorsagia circumcincta  
ORGANISM Teladorsagia circumcincta

REFERENCE Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongyloidea; Haemonchidae; Ostertagiinae; Teladorsagia.  
AUTHORS 1 (bases 1 to 800)  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guillian, D.,  
Hall, N., Quayle, M., and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Main Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCRFP1  
BACKWARD: T7PL  
Plate: 39 row: D column: 08  
Seq primer: TEXP1  
High quality sequence start: 35  
High quality sequence stop: 550.  
Location/Qualifiers

FEATURES

Source

1. 800  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="TC\_ad2\_39D08"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplex2; Site\_1: BamHI; Site\_2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match 61.1%; Score 466.2; DB 6; Length 800;  
Best Local Similarity 78.3%; Pred. No. 5.7e-130;  
Matches 574; Conservative 0; Mismatches 153; Indels 6; Gaps 1;

QY 19 GGCTGTGTAGTGTCTGTTCTCTGCGCCCTGTTCTCCTATGCGGAGCAGCCTTTGTTG 78  
DB 58 GGCTGTCAAGTGTGTTGTTCTCTGCTGCTGTGTCTCTCTATGCCAAGCAGGCTTTGTTG 117  
QY 79 TCCGAATAGTCTAAGCCAAAGTGAACGCGGAGCAGATTTTCTCGATTTTCACAATGA 138  
DB 118 CTCAAAACCTCTGGCCAAACCGCAACGAGCAGATTTTCTCGATTTTCACAATGA 177  
QY 139 TGTTCGTGAAATATAGCACTTGAATGTTGATTAAGTGAACAGTAATGACAGCGC 198  
DB 178 TATTGTCGGAACATAGCACTTGAAGAAGCTTGTGTAAGT-----TCACGACGATCC 231  
QY 199 GGTCAATCTTGTGCTCAGCTCAGAACATGTACAAAGTGAAGTGGAGTTCGAATGGAAGA 258  
DB 232 GGTGTTCTTGTGCTCAGCTCAAAACATGTACAAAGTGAAGTGGAGTTCGAATGGAAGA 291  
QY 259 AGTAGCAGCACAACAGATTGGCCCATGCAATGATCCCTACCGATTAATACCAAGCTGGC 318  
DB 292 AAAAGCAGCCAGCAGATTTGGCAATGACAGGTTCTCTTACCAATAGATCCCAAGCTGGC 351  
QY 319 TCAAAATATCGTAGATGGCTGTACTTCAAAAGCAGTGAAGAGAGACAGTTCTGCAACA 378  
DB 352 TCAAAATATCGCAAGATGGCTGTATTAATGCCAACAGTGAAGAGAGCAAGGTTTGAAGACA 411  
QY 379 AGTATCGTGTATTTGGGTGAGCGCATCGCTGGGATTTATGAAGCAGCAAACTTGACCA 438  
DB 412 AGTACCGTGTCTTGGGTGAGCGCATCGCTGGGATTTATGAAGGCAAGGCAAGCTTGATCG 471

QY	439	ATTGCTAACCAAGTGGGCTGGAACCTTAGCAAAACATTGCAAACTATAGAAACCGAAAGGT	498
Db	472	TTTGTCAACCAAGTGGGCTGGAACCGCTTGCCAAACATTGCTAACCTGGA AAAATCCGAAGGT	531
QY	499	TGGATGTGCCCATAGATCTGCCCCGCTCAGCAAAAACATGGTAGTATCCTGCGGTATGG	558
Db	532	TGGATGTGCCCTACAAAGATCTGCCCCGCTGAAAAACATGGTAGTATCCTGTGTGTATGG	591
QY	559	AAGCCCCAAACTTGCAACCGAAGAAAGTTATCTGGCAGGAAGAAAGCGTTGTGTGCGA	618
Db	592	AAGCCAAAACTTTCACCAACGAAGTAAATTGGAGCAAGGAAGTACTTGCAGTGCAA	651
QY	619	CGCTCGTCCAGATTCATTCTGCTGCGCAACAACCTGTGTGACAACGCAAGATGCTGCGAAGT	678
Db	652	CTCTTATCCGGATTGCTTTGTGCTGTGACAGCCTGTGTGACACGCATGAGCTGCGAGCT	711
QY	679	TCCGCCAACCAAGTGTGGCGCTGCGCCATGAAGCGAAAAAATGGTAGTCAACCCGAATAA	738
Db	712	TCGCTAACCAAGTGTGTGATGCAATAAGCGGAGAAAAAATGTGTAAACATCTTGAGCAA	771
QY	739	AATATTCATGCAA 751	
Db	772	AATATTAATTTAA 784	

RESULT 15	CB036762	LOCUS	DEFINITION
	CB036762	758 bp	mRNA
	Tc_ad2_25G07	linear	EST 15-JAN-2003
	_TEXT1	Teladorsagia circumcincta adults	library 2
	Teladorsagia circumcincta	cDNA clone TC ad2_25G07 5'	similar to
	CAD56659	Ancylostoma-secreted protein-like	prote
		Ostertagia,	mRNA
		sequence.	

ACCESSION	CB036762	
VERSION	CB036762.1	GI:2756007
KEYWORDS	EST.	
SOURCE	Teladorsagia circumcincta	
ORGANISM	Teladorsagia circumcincta	

REFERENCE	1 (bases 1 to 758)
AUTHORS	Blaxter, M.L., Parkinson, J., Whittton, C., Daub, J., Gulliano, D., Hall, N., Quayle, M. and Barrel, B.
TITLE	Edinburgh University/Sanger Centre Nematode EST project
JOURNAL	Unpublished (2000)
COMMENT	Contact: Blaxter ML

Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCRF1  
BACKWARD: T7PL  
Plate: 25 row: G column: 07  
Seq primer: TEXF1  
High quality sequence stop: 495.  
Location/Qualifiers  
1..758

```

1. . /50
/organism="Teladorsagia circumcincta"
/mol_type="mRNA"
/db_xref="taxon:45464"
/clone="Tc_ad2_25G07"
/sex="mixed"
/dev_stage="adult"
/clone_lib="Teladorsagia circumcincta adults library 2"

```

**ORIGIN**

/note="Vector: pTriplEx2; Site\_1: BamHI; Site\_2: BamHI; Teladorsagia circumcincta is a parasitic nematode. The library was constructed from mRNA from Teladorsagia circumcincta mixed adults."

Query Match	61.1%;	Score 466;	DB 6;	Length 758;
Best Local Similarity	79.5%;	pred. No. 6.5e-130;		
Matches 566;	Conservative	0;	Mismatches 140;	Indels 6;
				Gaps 1;

OY		19	G G C T G T T G A G T T G C T G T T C T C C T G G C C C T G T T C T C C T A T G C C G A A G C A G C C T T T G T T G	78
Db		25	G G C A G T C A C A G T T G G T G T T C T C T G G C T T T G T C C T C C A T G C C G A A G C A G C C T T T G T T G	84
OY		79	T C C G A A T A G T C T A A G C C C A A G T G A C A G C G C A G G C A G A T T T C T C G A T T T T C A C A A T G A	138
Db		85	C C C A A A A A C T C T A A G C C A A C T G A C A A C A C G A G C A G A T T T C T C G A T T T C A C A A T G A	144
OY		139	T G T T C G T C G A A T A T A G C A C T T G G A A T G T T G A T A A C T G C A C A G T A A A T G C A G A C G C	198
Db		145	T A T T C G T C G G A A C A T A G C A C T C G G A A A G A G C T T G T G A A C T -----T C A C G A C T G A T C C	198
OY		199	G G T C A T T C T T G G T C C A G C T C A G A C A T G T A C A A A G T G A C T G G G A T T G C A C T T G G A A G A	258
Db		199	G G T T G T T C T T G G T C C A G C T C A G A C A T G T A C A A G C T G G A C T G G G A C T T G A C C T T G G A A C A	258
OY		259	A G T A G C A G C A C A C A G A T T G C G C C A T G C A A T G A T C C C C T A C C G A T A A T A C C A G C C T G G C	318
Db		259	A A A A G C A G C T C A A C A G A T T G C G A A T G C A C G G T T C C T T A C C A A T A G A T C C C A G C C T G G C	318
OY		319	T C A A A A T A T C G T A G A T G G C T G T A C T T C A A A G A C A G T G A A G A A G A C A G T T C T G C A A C A	378
Db		319	T C A A A A T A T C G C A A G A T G G C T G T A T T A T T C C A A C A G T G A A G A A G A C A G G T T T G A G A C A	378
OY		379	A G T A T C G T G T A T T G G G T G A G C G C A T C G T G G A T T A T G A A A G G C A C G A A C T T G A C C A	438
Db		379	A G T A C C G T G T C T T G G G T A C C C C A T C G C T A C G A T T C A T G A A G G G C A C G G C A C T T G A T C G	438
OY		439	A T T T G C T A A C C A G T G G G C T G A A C C T C T A G C A A A C A T T G C A A C T A T A G A A C C G A A A G G T	498
Db		439	T T T C G C T P A A C C A G T G G G C T G A A C C G C C T T G C C A C A C A T T G C T A C T G A A A A A T C G G A A G G T	498
OY		499	T G G A T G T C C C A T A A G A T C T G C C C C G C T C A G C A A A A C A T G T A G T A T C C T G C G T A T G G	558
Db		499	T G G A T G T C C T T A C A A G A T C T G C C C C G C C T G A A A A A C A T G T C G T A T C C T G T G T A T G G	558
OY		559	A A G C C C C A A A C T T G C A C C G A A C G A A G T A T A T C T G G C A G A A G G A A A G G C T T G T G T G C G A	618
Db		559	A A G C C A A A A A C T T T C A C C A A C G A A G T A A T T T G G A G C A A G A A G T A C T T G C A G T G C A A	618
OY		619	C G C T C G T C C A G A T T C A T T C T G C T G C G A C A A C C T G T G T G A C A C G C G A G A T G C T G C G A G T G T	678
Db		619	C T C T T A T C C G G A A T C G T T T T G C T G T G A C A C C C T G T G T G A C A C G C G T G A G C T G C G A G C C T	678
OY		679	T C G C C A C C A G T G T T G C G C G T C G C C A T G A A G C G A A A A G A A T T G T A G T A C C	730
Db		679	T C G C A C C A G T G T T G C T G A T C G A C A T G A G C C G A G A A G A A T T C T G T T A C A T C	730

Search completed: March 4, 2006, 08:18:32  
Job time : 3919 secs

Search completed: March 4, 2006, 08:18:32  
Job time : 3919 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:02:37 ; Search time 229 Seconds  
(without alignments)  
711.690 Million cell updates/sec

Title: US-10-527-771-10  
Perfect score: 1253  
Sequence: 1 MSAVVVAVLALFSTYAEAG.....DNLCTRDASVRHQCASP 231

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1253	100.0	231	2	Q8IFT6 OSTOS
2	638.5	51.0	236	2	Q8T8F4 OSTOS
3	277	22.1	424	1	ASP_ANCCA
4	274	21.9	425	2	Q816X0 9BILA
5	274	21.9	425	2	Q77153 9BILA
6	263	21.0	424	2	Q76744 NECAM
7	260	20.8	424	2	Q9XZ41 ANCCA
8	211	16.8	424	2	Q8MQ79 CABEL
9	211	16.8	425	2	Q19348 CABEL
10	205	16.4	429	2	Q614N2 CABBR
11	195.5	15.6	220	2	Q602Y9 CABBR
12	191	15.2	424	2	Q86GK5 ANCCA
13	186.5	14.9	451	2	Q86GK4 ANCCA
14	178.5	14.2	221	2	P90959 CABEL
15	176	14.0	248	2	Q9BIQ8 9BILA
16	167	13.3	196	2	Q62507 CABEL
17	166	13.2	217	2	Q7Z1H0 9BILA
18	164	13.1	248	2	Q9BIQ6 9BILA
19	161.5	12.9	217	2	Q7Z1H2 9BILA
20	160	12.8	241	2	Q76663 CABEL
21	158.5	12.6	208	2	Q9N5N3 CABEL
22	156	12.5	210	2	Q7Z1H1 NECAM
23	152.5	12.2	218	2	Q77221 ANCCA
24	151	12.1	465	2	Q9BIP2 9BILA
25	151	12.1	491	2	Q9BIQ4 9BILA
26	150	12.0	231	2	Q9BIQ7 9BILA
27	146	11.7	212	2	Q62508 CABEL
28	142.5	11.4	207	2	Q93746 CABEL
29	142.5	11.4	211	2	Q625N2 CABBR
30	138	11.0	200	2	Q86GK7 ANCCA
31	135.5	10.8	213	2	Q93747 CABEL

32	135	10.8	246	2	Q9N3I3 CABEL	Q9n3i3 caenorhabdi
33	135	10.8	459	2	Q45132 HABCO	Q45132 haemochus
34	132.5	10.6	208	2	Q9N5N4 CABEL	Q9n5n4 caenorhabdi
35	132.5	10.6	212	2	Q61FT8 CABBR	Q61ft8 caenorhabdi
36	131.5	10.5	224	2	Q86FX1 ONCVO	Q86fx1 onchocerca
37	130.5	10.4	251	2	Q5FVN7 RAT	Q5fvn7 rattus norv
38	130	10.4	209	2	Q623I7 CABBR	Q623i7 caenorhabdi
39	129	10.3	206	2	Q623H6 CABBR	Q623h6 caenorhabdi
40	128.5	10.3	220	2	Q16854 ONCVO	Q16854 onchocerca
41	127.5	10.2	361	2	Q624A6 CABBR	Q624a6 caenorhabdi
42	125.5	10.0	213	2	Q60KM5 CABBR	Q60kms caenorhabdi
43	124.5	9.9	224	2	Q9GPN4 ONCVO	Q9gpn4 onchocerca
44	124.5	9.9	303	2	Q7Z4U9 HUMAN	Q7z4u9 homo sapien
45	124.5	9.9	446	2	Q6DCB3 HUMAN	Q6dcb3 homo sapien

ALIGNMENTS

RESULT 1									
ID	Q8IFT6 OSTOS	Q8IFT6 OSTOS PRELIMINARY;	PRT;	231	AA.				
AC	Q8IFT6;								
DT	01-MAR-2003 (Tremblrel. 23, Created)								
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)								
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)								
DE	Ancyclostoma-secreted protein-like protein.								
GN	Name=aasp2;								
OS	Ostertagia ostertagi.								
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;								
OC	Trichostrongyloidea; Haemonchidae; Ostertaginae; Ostertagia.								
OX	NCBI_TaxID=6317;								
RN	[1]								
RP	NCU00100.1 SEQUENCE.								
RX	MEDLINE=22593580; PubMed=12706806; DOI=10.1016/S0166-6851(03)00044-6;								
RA	Geldhof P.B., Vercouteren I., Gevaert K., Staes A., Knox D.,								
RA	Vandekerckhove J., Vercruyse J., Claerebout B.;								
RT	"Activation-associated secreted proteins are the most abundant								
RT	antigens in a host protective fraction from Ostertagia ostertagi. "								
RL	Mol. Biochem. Parasitol. 128:111-114(2003).								
DR	EMBL; AJ515523; CAD56659.1; -; mRNA.								
DR	GO; GO:0005576; C:extracellular region; IEA.								
DR	InterPro; IPR001283; Allrgn_V5/Tpkl.								
DR	Pfam; PF00188; SCP; 1.								
DR	SMART; SM00198; SCP; 1.								
DR	SEQUENCE 231 AA; 25504 MW; 8E1043E42AF07081 CRC64;								
Query Match									
Best Local Similarity 100.0%; Score 1253; DB 2; Length 231;									
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1 MSAVVVAVLALFSTYAEAGFCCPNSLSQSDSARQIFLDFHNDVRNIALGNGLINWTVN	60							
DB	1 MSAVVVAVLALFSTYAEAGFCCPNSLSQSDSARQIFLDFHNDVRNIALGNGLINWTVN	60							
QY	61 ADAVITLGAQNMVYKVDWDCNLEBVAQAQIAPCNDPLPINTSLAQNIARWLYFKDSEETV	120							
DB	61 ADAVITLGAQNMVYKVDWDCNLEBVAQAQIAPCNDPLPINTSLAQNIARWLYFKDSEETV	120							
QY	121 LQQVSWYVWSASLGFNKGTKLDQFANQWAEPLNIANYRNKVGCAHKICPAQNMVWSC	180							
DB	121 LQQVSWYVWSASLGFNKGTKLDQFANQWAEPLNIANYRNKVGCAHKICPAQNMVWSC	180							
QY	181 VYGSPKLAPEVIMWEGKACVCDARPDFCCDNLCTRDASVRHQCASP 231								
DB	181 VYGSPKLAPEVIMWEGKACVCDARPDFCCDNLCTRDASVRHQCASP 231								
RESULT 2									
ID	Q8T8F4 OSTOS	Q8T8F4 OSTOS PRELIMINARY;	PRT;	236	AA.				
AC	Q8T8F4;								
DT	01-JUN-2002 (Tremblrel. 21, Created)								



```
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
OS Ancylostoma-secreted protein-like protein (Fragment).
SN Name=OSP-2;
OC Oostertagia oostertagi.
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Ostertaglinae; Ostertagia.
ON NCBI_TaxID=6317;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22503448; PubMed=12615319; DOI=10.1016/S0166-6851(02)00274-8;
RA Vercauteren I.J.R., Geldhof P., Peelaerts I., Claerebout B., Berx G.,
RA /ercruyse J.;
RT "Identification of excretory-secretory products of larval and adult
RT Ostertagia oostertagi by immunoscreening of cDNA libraries.";
RL Mol. Biochem. Parasitol. 126:201-208(2003).
DR EMBL; AJ310812; CAD23183.1; -; mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR InterPro; IPRO01283; Allrgn_V5/Tpx1.
DR Pfam; PF00188; SCP; 1.
DR SMART; SMO0198; SCP; 1.
FT NON TER 1 1
SQ SEQUENCE 236 AA; 26228 MW; DB52781B8353A698 CRC64;

Query Match          51.0%; Score 638.5; DB 2; Length 236;
Best Local Similarity 52.5%; Pred. No. 2e-49;
Matches .125; Conservative 36; Mismatches 66; Indels 11; Gaps 4;

QY      1 MSAVVAVVAV-LTALFSYAEGFCCPNSLSQSDSARQIFLDHNDVRNIALGGLINMT 58
        |||::||::||::|||||::||::|||::|||::|||::|||::|||::|||
DB       1 MGALIGIAALYLVLVTSNTBAGFCPCPADLNQTDEARXIFLDFHNQVRDIDAGASPLNLT 60

QY      59 VNADAV----ILGPAQNMYKYWDNCNLBEVAQAQIAPCNDPLPINTSLAQNIARWLYFKD 114
        ||::||::||::||::|||::|||::|||::|||::|||::|||::|||::||
DB       61 ---GAQMRRNVLGPAKNMTRMDWCNLEAKAKAMTWCTPLPIDTSPONTAQWLFFQN 117

QY      115 SEETVLOQVSWMYWSASLGFMKGTQLDPANQMAEPLANIANRYNRKVGCAHKIC--PA 172
        :|:|||:|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       118 SEXEVTLTOTPMWSWTASLRNLQPDTBANINYMQIRPLSIANWMLKVGCCHKVKPEPT 177

QY      173 QGNMVSVCGYSPIKLAPNEVIWQEGKACVCDDARPDSFCEDNLCDTRDAASVRHQCCAS 230
        |||||:|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       178 GTNWVTSCAYGGBVLQDNVEVVDKGPTCMCNAYPNISFCNNLCDTIAAATLRNQPKS 235

RESULT 3
ASP_ANCCA
ID ASP ANCCA STANDARD; PRT; 424 AA.
AC Q16937;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ancylostoma secreted protein precursor.
GN Name=ASP;
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiada; Strongyliada;
OC Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96215086; PubMed=8636085; DOI=10.1074/jbc.271.12.6672;
RA Hawdon J.M., Jones B.F., Hoffman D.R., Hotez P.J.;
RT "Cloning and characterization of Ancylostoma-secreted protein. A novel
RT protein associated with the transition to parasitism by infective
RT hookworm larvae.";
RL J. Biol. Chem. 271:6672-6678(1996).
CC -!- FUNCTION: Associated with the transition to parasitism by
CC infective hookworm larvae.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: Belongs to the CRISP family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
```

[illegible]

	Query Match	21.9%	Score 274;	DB 2;	Length 425;	
	Best Local Similarity	30.5%;	Pred. No. 3.5e-16;			
	Matches	74;	Conservative	34;	Mismatches 101;	Indels 34; Gaps 9
Qy	2	SAAVVAVLAL---	FSYAEAGFCCPNLSIQSDSAROIFLDFHNDVRNIALG----	N	52	
		: : : :	:  :                    :  :  :	:		
Dd	3	SPVVVISVLTVAFCDASPVKASFCSNS-GITDSDRQAFLDFHNARRVAQGVEDNK	S	61		
Qy	53	GLINWTVNADAVILGPAAQNMYKVMDNCULEEVAACQIAPCNDPLPINTSLAQNIARWL--		110		
		:  :	:      :  :  :	:		
Dd	62	GKLN-----PAKNMKLDWDCEMEQKLQDAIQSCPGFAGIGVAQNIISWSGS		110		
Qy	111	-YFKDSEETVLQOVSMYWVSASLGFMKGTKLDQFANOWAEPPLANIANRYNRKKVGCAHKI		169		
		:  :	:  :  :  :  :  :	:		
Dd	111	GGFPNPSEKINSTLASWMGGAQNNGVASDNK---YTGGGLYAFASNMVFSETTKLGCAYKV		167		
Qy	170	CPAQQNMVWSCVYGSPKLAPNEVIWGEGKACV----CDARPDSFCDCNUJCDT-RDAASVR		224		
		:  :  :  :  :	:  :  :  :  :  :	:		
Dd	168	CGTK--LTLSCTYNGIGYMTGAPMWETGOACKAKAGADCTTFPKNSGCCEDGLCTKGADVPETN		225		
Qy	225	HQC	227			
Dd	226	QOC	228			

```

RESULT 5
077153_9BILA
ID 077153_9BILA PRELIMINARY; PRT; 425 AA.
AC 077153;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Ancylostoma-secreted protein 1 precursor.
GN Name=aspl1;
OS Ancylostoma duodenale.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdiltida; Strongylida;
OC Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=51022;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Shanghai;
RX MEDLINE=96215086; PubMed=8636085; DOI=10.1074/jbc.271.12.6672;
RA Hawdon J.M., Jones B.F., Hoffman D.R., Hotez P.J.;
RT "Cloning and characterization of Ancylostoma-secreted protein. A novel
RT protein associated with the transition to parasitism by infective
RT hookworm larvae.";
RL J. Biol. Chem. 271:6672-6678(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Shanghai;
RX MEDLINE=99151774;
RA Bin Z., Hawdon J., Qiang S., Hatnan R., Huifang Q., Wei H.,
RA Shu-hua X., Tlehua L., Xing G., Zheng F., Hotez P.;
RT "Ancylostoma secreted protein 1 (ASP-1) homologues in human
RT hookworms.";
RL Nml. Biochem. Parasitol. 98:143-149(1999).
DR EMBL; AF077402; AAD13339.1; -; mRNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR InterPro; IPR001283; Align_V5/Tp1.
DR InterPro; IPR002413; V5_allergen.
DR Pfam; PF00188; SCP_2.
DR PRINTS; PRO0838; VSALLERGEN.
DR PRINTS; PRO0837; V5TPXLIKE.
DR SMART; SM00198; SCP_2.
DR PROSITE; PS01009; CRISP_1; 1.
DR Signal.
FT SIGNAL. 1 19 Potential.
FT CHAIN 20 425 ancylostoma-secreted protein 1.
FT SEQUENCE 425 AA; 45821 MW; 1F6B9D7E62EBDA8 CRC64;
Query Match 21.9%; Score 274; DB 2; Length 425;
Best Local Similarity 30.3%; Pred. No. 3.5e-16;

```

	Matches	74;	Conservative	37;	Mismatches	97;	Indels	36;	Gaps	11;			
QY	2	SA	AAVVAVLLAL----	PS	YAEGFCCPNSLSQSDSARQIFLDFHNDVRRNIALG-----	N	52						
		:	:	:	:	:	:	:	:	:			
Db	3	SS	VVISVISTIAFC	DA	SPARASFGCSNN-GITDSRQAF	LD	FHNNARRRVAQGVEDNKS	61					
QY	53	GL	INMTVNADAVIL	GP	AQNNYKVDWDCNLE	EVAAQQA	IAPCNDPLPINTSLAQNIA	WLY-	111				
		:	:	:	:	:	:	:	:	:			
Db	62	GK	LN-----	PA	KNYKLEWDCRMEQQLD	AIQSCPGSAGIQGFSQNVMS	MSNS	110					
QY	112	--	FKDSEETVLQOV	SW	TVWSASLGFMK	GTKLD-QP	ANQWAEPLANIAN	YRNRRKVGCAHK	168				
		:	:	:	:	:	:	:	:	:			
Db	111	GG	FNSSEK-IE	ST	LSGWSGAK--	NN	GVSDNKYTGGLYAF	SNMVFSETTKIG	CAYK	166			
QY	169	IC	PAQQNMVSCV	YGS	PXLA	PN	EVIMQEGKACV----	CD	ARPD	SFCCDNLCDT-RD	ASV	223	
		:	:	:	:	:	:	:	:	:	:		
Db	167	VG	GTK--	MAT	SCIYNGIS	YITNA	PMMETGQAC	KTAGD	CSTYKN	SGCEDSLCTK	GADV	PET	224
QY	224	RH	QC	227									
		:	:										
Db	225	NO	QC	228									

Query Match	Best Local Similarity	Score	DB	Length
Matches 73; Conservative	30.5%;	263;	2;	424;
	36; Mismatches	106;	Indels	24; Gaps

```
QY      1 MSAAVVAVLLAL-----FSYAAGFCPCPSLSQSDSARQIFLDFHNDVRNIALGNGLIN 56
       |||::: | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MSPPVVSVFETIAFCNASPARDSFGCSNS-GITDSDROAFLDFHNARRRVAKG----- 54

QY      57 WTVNADAVILGPAQNMYKVDWDCNLBEVAAQIAPCNDPLPINTSLAQNIARWLFEKDSE 116
       :::: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      55 -LEDSNSGKLNPAKNMKYKLSWDCAEQQLQDAIQSCPSGFAGIQGVAAQNTMSSSSGGYP 113

QY      117 EETVLQQ--VSWYWVASLGFMTKTL-D-QFANQWAEPLANIANRYRNRYGCAHKICPAQ 173
       ::: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      114 DPSVKIEPTLSGWMSGAK--XNGVGPDNKYTGGLFAFSNMVYSSETTKLGCAVKCGTK 170

QY      174 QNVVSCVYYGSFKLAIPNEVIWQEGKACV----CDARPDSFCCDNLCDDT-RDAASVRHQ 227
       ::|||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      171 --LAVSCTYNGVGYITNQPMMETGQAQCQTGADCSITYKNSGCEDGLCTKGPDPVEPTNQOC 227
```

RESULT	7
ID	09XZ41_ANCCA
ID	09XZ41_ANCCA PRELIMINARY; PRT; 424 AA.
AC	Q9XZ41;
DT	01-NOV-1999 (TREMBLrel. 12, Created)
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE	Ancylostoma-secreted protein 1 precursor.
OS	Ancylostoma caninum (Dog hookworm).
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC	Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX	NCBI_TaxID=29170;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=Shanghai;
RX	MEDLINE=20163524; PubMed=10701589;
RA	Shan Q., Zhan B., Xiao S.-H., Feng Z., Hotez P., Hawdon J.M.;
RT	"Variation between ASP-1 molecules from Ancylostoma caninum in China and the United States."
RL	J. Parasitol. 86:181-185 (2000).
DR	EMBL; AF132291; AAD31839.1; -; mRNA.
DR	HSSP; P04284; 1CFE.
DR	GO; GO:0005576; C:extracellular region; IEA.
DR	InterPro; IPR01283; Allrgn_V5/Tpxl.
DR	InterPro; IPR002413; V5_allergen.
DR	Pfam; PF00188; SCP; 2.
DR	PRINTS; PRO0838; VSALLERGEN.
DR	PRINTS; PR00837; V5TPXLKE.
DR	SMART; SM00198; SCP; 2.
DR	PROSITE; PS01009; CRISP_1; 1.
KW	Signal.
FT	SIGNAL. 1 18 Potential.
FT	CHAIN 19 424 ancylostoma-secreted protein 1.
FT	SEQUENCE 424 AA; 45761 MW; 8409CDF8AECD248E CRC64;
Query Match	20.8%; Score 260; DB 2; Length 424;
Best Local Similarity	30.5%; Pred. No. 6.4e-15;
Matches	75; Conservative 33; Mismatches 100; Indels 38; Gaps 11,
OY	1 MSAAVVAVLLAL-----FSYAEGFCCPNLSQSDSARQIFLDHFNDVRNIALGNGLIN 56
DB	1 MFSPVVSVFITTAICDASPARDSEFGCSNS-GITDKDQAFLDFHNARRRVAKG----- 54
OY	57 WTVNADAVILGPAQNMYKVDWDNCLEEVAAQOIAPCNDPLPINTSLAQNIARWL---FK 113
DB	55 -LEDNSNGKLNPANMYYKLSWDCAMEQQQLDPAIQSCPSGFAGIQQVNAQNTMSWSSGGFP 113
OY	114 DSEBETVLQOVSWYVVSASLGFMKGTCLDQ-----FANOAEPLANIANRYNRKVGCA 166
DB	114 DPSVK-IEPTLSGWSGAK--KNGVGPDNKYNGGGLFA-----FSNNVYSETTKLGCA 163
OY	167 HKICPAQQNMYSGCYGSPKLAPNEVIWOEGKACV----CDARPDSFCCDNLCDT-RDA 221
DB	164 YKVCGTK--LAVSGTYNGVGIITNQPMMETGAQCTGADCDSTYKNSGCEDGLCTKGPDPV 221
OY	222 SVRHQC 227

Db 222 ETNQC 227

RESULT 8					
ID	Q8MQ79_CABEL	PRELIMINARY;	PRT;	424 AA.	
AC	Q8MQ79;				
DT	01-OCT-2002 (TReMBLrel. 22, Created)				
DT	01-OCT-2002 (TReMBLrel. 22, Last sequence update)				
DT	01-JUN-2003 (TReMBLrel. 24, Last annotation update)				
DE	Venom-allergen-like protein protein 1, Isoform b.				
DN	Name=vap-1; ORFNames=FlIC7.3;				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;				
CC	Rhabditiidae; Peloderinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].				
RC	STRAIN=Bristol N2;				
RX	MEDLINE=99069613; PubMed=9851916;				
RG	The C. elegans sequencing consortium;				
RT	"Genome sequence of the nematode C. elegans: a platform for investigating biology.";				
RL	Science 282:2012-2018(1998).				
DR	EMBL; U42839; AAM54195.1; -; Genomic_DNA.				
DR	HSSP; Q05110; IQNX.				
DR	Ensembl; FlIC7.3; Caenorhabditis elegans.				
DR	WormBase; WBGene0006886; vap-1.				
DR	WormPeP; FlIC7.3b; CE30944.				
DR	GO; GO:0005576; C:extracellular region; IEA.				
DR	InterPro; IPR001283; Allrgn_V5/Tpxl.				
DR	InterPro; IPR002413; V5_allergen.				
DR	Pfam; PF00188; SCP_2.				
DR	PRINTS; PRO0838; VSALLERGEN.				
DR	PRINTS; PR00837; V5TPXLKE.				
DR	SMART; SM00198; SCP_2.				
DR	PROSITE; PS01009; CRISP_1; 1.				
KW	Complete proteome.				
SQ	SEQUENCE 424 AA; 46224 MW; 400AA4257F78B4ACB CRC64;				
Query Match 16.8%; Score 211; DB 2; Length 424;					
Best Local Similarity 26.5%; Pred.No.1.7e-10;					
Matches 65; Conservative 36; Mismatches 76; Indels 68; Gaps 10.					
QY	3 AAVVAVLLALFSYAEB-GFCCPNLSQSDSARQIFLDFHNDVRNIAG---- <td>56</td> <td></td> <td></td> <td></td>	56			
	:   :           :   :   :   :   :   :   :				
Db	2 AVLAVVLLLACLERAQAQTFCGSGNT-KINDQARKMFYDAHNDARSMAKGLSPNKCGL--	58			
QY	57 WTVNADAVILGPQNMYKVWDPCNLBEVAQAQIAPC-----ND	94			
	:   :   :   :   :   :   :   :   :   :   :				
Db	59 -----LSGGKNVELNMDCENBAKAQEWAEGCPSSFQTFDPPTWGNYATYMGSIAD	109			
QY	95 PLPIINTSLAQNIARWLIFYKDSSEETVLQQVSWMYVWSASLGFMKGITLDQPANQWAEPLAN	154			
	:   :   :   :   :   :   :   :   :   :				
Db	110 PLPY-ASMAYN-----GWWEIRTVGLTDPPD-NKYTN SAMFRPAN	147			
QY	155 IANYRNRYCAAHKICPAQQNNMVSGSPKLAPNEVIWOEGKACVCDAE---PDSFC	210			
	:   :   :   :   :   :   :   :   :   :   :				
Db	148 MANGKASAFGCAYALCAGK-LSINCINIKIGYMTNAIIYEKGDACTSDAECTTYSDSQ	205			
QY	211 CDNLC 215				
	:   :				
Db	206 KNGLC 210				
RESULT 9					
ID	Q19348_CABEL	PRELIMINARY;	PRT;	425 AA.	
AC	Q19348;				
DT	01-NOV-1996 (TReMBLrel. 01, Created)				
DT	01-JUN-1998 (TReMBLrel. 06, Last sequence update)				
DT	13-SEP-2005 (TReMBLrel. 31, Last annotation update)				



```

DE Venom-allergen-like protein protein 1, isoform a.
OS Name=vap-1; ORFNames=F11C7.3;
GN Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=N2 Bristol;
RL Liu L.X.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U42839; AAC69015.1; -; Genomic_DNA.
DR EMBL; AF112356; AAD27559.1; -; mRNA.
DR PIR; C89753; C89753.
DR HSSP; Q05110; IQNX.
DR Ensembl; F11C7.3; Caenorhabditis elegans.
DR WormBase; WBGene0006886; vap-1.
DR WormPep; F11C7.3a; CE09350.
DR GO; GO:0005576; C:extracellular region; IEA.
DR InterPro; IPR001283; Allrgn_V5/Tpx1.
DR InterPro; IPR002413; V5_allergen.
DR Pfam; PF00188; SCP_2.
DR PRINTS; PRO0838; VSALLERGEN.
DR PRINTS; PRO0837; V5TPXLIKE.
DR ProDom; PD000542; Allrgn_V5/Tpx1_1.
DR SMART; SM00198; SCP_2.
DR PROSITE; PS01009; CRISP_1; 1.
DR Complete proteome.
SQ SEQUENCE 425 AA; 46323 MW; 9B7233B8A7340EEB CRC64;

Query Match 16.8%; Score 211; DB 2; Length 425;
Best Local Similarity 26.5%; Pred. No. 1.7e-10;
Matches 65; Conservative 36; Mismatches 76; Indels 68; Gaps 10;

QY 3 AAVVAVLLALFSYAEA-GFCCPNSLSQSDSARQIFLDFHNDVRNIALG-----NGLIN 56
   | : | : ||| | | | | | : | : | : | | | | | | | | | |
DB 2 AVLA VVLLLA CLERAVAQTFCGCSNT-KINDQARKMFYDAHNDARRSMAKGLEPNKCGL-- 58

QY 57 WTVNADAVILGPAQNMKYKVDWDCNLEBVAQOIAPC-----ND 94
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 59 -----LSGKKNVIELNWDCMEAKAQEWADGCPSSFQTFDPTWGQNYATYMGSIAD 109

QY 95 PLPINTSLAQNIARWLTFKDSERTVLQOVSWYVWSASLGFMKGTLDQFANQWABPLAN 154
   ||| | | | | | | | | | | | | | | | | | | | | | | |
DB 110 PLPY-ASMAVN-----GWMSEIRTVGLTDPD--NKYTNSAMFRFAN 147

QY 155 IANYRNKVGCAHKICPAQQNMVVS CVYGS PKLAPNEVIMQEGKACVCDAR----PDSFC 210
   : || : ||| : | : : : | : : : | | | | | | | | | |
DB 148 MANGKASAFGCAYALCAGK--LSINCINNKIGYMTNATIIYKGDAC TSDAECTTYSDSQC 205

QY 211 CDNLC 215
   : ||
DB 206 KNGLC 210

RESULT 10
Q614N2 CAEBR PRELIMINARY; PRT; 429 AA.
AC Q614N2;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein CBG15925.
GN Name=CBG15925;
OS Caenorhabditis briggsae.

```

```

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC0100076; CAB9675.1; -; Genomic DNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR InterPro; IPR001283; Allrgn_V5/Tpx1.
DR InterPro; IPR002413; V5_allergen.
DR Pfam; PF00188; SCP; 2.
DR PRINTS; PR00838; V5ALLERGEN.
DR PRINTS; PR00837; V5TPXLIKE.
DR ProDom; PD000542; Allrgn_V5/Tpx1; 1.
DR SMART; SM00198; SCP; 1.
DR PROSITE; PS01009; CRISP_1; 1.
KM Hypothetical protein.
SQ SEQUENCE 429 AA; 46395 MW; 427C8AC26A3AE091 CRC64;

Query Match 16.4%; Score 205; DB 2; Length 429;
Best Local Similarity 29.4%; Pred. No. 6.1e-10;
Matches 68; Conservative 31; Mismatches 100; Indels 32; Gaps 9;

QY 1 MSAVVVAVLALFSYAAGFCCPNSLSQSSARQIFLDFHNDVRRLALGNGLINVTN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MAFVAVLVLAACLEGTFPAQFGCTNAKS-SDRAKMFQDAHNDARRSWAKG-----LEP 53

QY 61 ADAVILGPAQNMVK-VDWDCNLBRVAAQIAPC-----NDPLPINTSLAQNIRWLYFK 113
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 54 NKGMLSEGNVYELVRWDCMEBAKAQTWADGCPSGFQSSDP-----TYGQNTMFTSGTF 108

QY 114 DSEBETVLQGVSNYVWSASLGFM-----KGTKLDQFANQWAEPLANTANYRNKKVGAHK 168
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 109 ADPVATAASAVNSWMAQVRSGGLTDPDNKXTSSSIFA--FSNVRETMANGKATAIGCAYA 166

QY 169 ICPAQQNMVSCVYSGPKLAPNEVIMQEGKACV---CDARPDSECCDNLC 215
   || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 167 ICGT-TLSVNCILNKIGYMTNAILYKGTACANSDCTTYADSECRNGLC 215

RESULT 11
Q60ZY9 CAEBR PRELIMINARY; PRT; 220 AA.
ID Q60ZY9; AC Q60ZY9;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein CBG17591.
DN Name=CBG17591;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC0100084; CAB70820.1; -; Genomic DNA.
DR GO; GO:0005576; C:extracellular region; IEA.
DR InterPro; IPR001283; Allrgn_V5/Tpx1.
DR Pfam; PF00188; SCP; 1.
DR SMART; SM00198; SCP; 1.
KM Hypothetical protein.
SQ SEQUENCE 220 AA; 25472 MW; 73D9CD577410429B CRC64;

Query Match 15.6%; Score 195.5; DB 2; Length 220;

```



DR pEam, PF00188; SCP, 1.  
DR PRINTS; PRO0838; VALLERGEN.  
DR SMART; SM00198; SCP, 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 221 AA; 25494 MW; FE2E53284DB06F89 CRC64;

Query Match	14.2%;	Score 178.5;	DB 2;	Length 221;
Best Local Similarity	29.3%;	Pred. No. 7.3e-08;		
Matches	51;	Conservative	35;	Mismatches 59;
			Indels	29;
			Gaps	8;

[illegible]

## RESULT 15

ID	Q9BIQ8_9BILA	PRELIMINARY;	PRT;	248 AA.
AC	Q9BIQ8;			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)		
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)		
DE	Activation associated secreted protein-like protein (Fragment).			
OS	Cooperia punctata.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;			
OC	Trichostrongylidae; Cooperiidae; Cooperia.			
OX	NCBI_TaxID=96640;			

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Cp-ASPvar-7;  
RX PubMed=12443689; DOI=10.1016/S0034-5288(02)00125-X;  
RA Yatsuda A.P., Eysker M., Vieira-Bressan M.C., De Vries B.,  
RT "A family of activation associated secreted protein (ASP) homologues  
of *Cooperia punctata*."  
RL Res. Vet. Sci. 73:297-306(2002).  
DR EMBL; AF352698; AAK35183.1; -; mRNA.  
DR HSSP; Q05110; IQNX.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR InterPro; IPR001283; Allrgn V5/Tpx1.  
DR InterPro; IPR002413; V5\_allergen.  
DR pfam; PF00188; SCP; 1.  
DR PRINTS; PR00838; VSALLERGEN.  
DR PRINTS; PR00837; V5TPXLIKE.  
DR SMART; SM00198; SCP; 1.  
DR PROSITE; PS01009; CRISP\_1; 1.  
FT NON TER 1 1  
SQ SEQUENCE 248 AA; 27558 MW; A7EDA83B9BEF7481 CRC64;

Query Match	14.0%;	Score 176;	DB 2;	Length 248;
Best Local Similarity	25.5%;	Pred. No. 1.4e-07;		
Matches	53;	Conservative	27;	Mismatches 100; Indels 28; Gaps 7;

```

Qy      23 CPNLSOSDSARQIFLDFHNDVRNRIALGNGLINWTVNADAVILGPAQNMYKYDMDCNLE 82
      |  :  |||: |||:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      49 CTLDNGMTDEARQVFLDKHNEYRQLVARGE-----AQNKTLGAPPAARMKLKRYDCDLB 102
Qy      83 EVAAQOIAPC---NDPLPINTSLAQNIARWLYF--KDSEETVLQOVSWYVWSASLGFM 136
      :  :  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      103 AHVMEHVAKCKGHSPPDVLKGRGQNI--WAITVPLDKAEAKRSVDDMYFHELTYYGIT 160
Qy      137 KGTKLDQFANQWAEPLANIANY-----RNRKVGCAHKI CPAQONMVVSCYVGSBKLA PN 190
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      161 ADNKIS-----IDNAKTI GHYSQVVMQKSNRLGCAAVSCPEQRRLYVGCETYPGNTLR 214
Qy      191 EVIWOEGKACVCDARPDSCCDNLCDTR 218

```

```

Db      215 HLTYDIGPCKRD--EDCKSSCRSTQ 240
          ::|: | : | | | :

```

Search completed: March 3, 2006, 19:09:28  
Job time : 231 secs



***This Page Blank (uspto)***

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

- Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).
- Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**This Page Blank (uspto)**



## **Protein Sequence Searches - February 2005**

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

***This Page Blank (uspto)***